



AI-READY ENTERPRISE PLATFORM FOR HIGHER EDUCATION



AI Accelerates the Reinvention of Higher Education with an End-to-End AI Platform from NVIDIA

Artificial intelligence is at the heart of the digital transformation that's reinventing higher education, driving demand for new research, curriculum development, and improved administrative services. Most universities know they need to invest in AI infrastructure but struggle to identify a cohesive strategy and platform that will enable sustainable, consistent success that's also aligned with their strategic goals.

Outside of AI, universities continue to face infrastructure challenges. As the need to support remote and hybrid approaches to learning grows, a university's compute, bandwidth, and storage needs increase proportionately. Enabling remote research and student development labs with compute-intensive applications puts tremendous pressure on already overburdened university IT infrastructure.

These existing IT infrastructure stressors exacerbate the difficulty of building the right compute systems to support university AI goals. Unlike other educational technology (edtech) applications, AI applications are a relatively recent development that tend to be configured on a project-by-project basis. They're anchored in rapidly evolving, open-source code and lack proven approaches that meet the rigors of scaled production settings. With competing priorities and a lack of visibility into the holistic needs for AI infrastructure across projects and learning environments, educational IT departments are pressed to create the environment their universities need to expand support for new AI use cases. These challenges may seem daunting, but two-thirds of institutions now see investments in AI as strategic. A majority of institutions have already started to experiment with AI as part of their strategy, and 38 percent have adopted it as a core part of the strategy.

Data centers in higher education are built to provide resilient, always-on services across students, faculty, administration, and researchers within the university. However, new technologies challenge the status quo many educational institutions have settled on.

- > **Existing traditional applications leveraged for virtual labs:** These include computer-aided design (CAD) and computer-aided engineering (CAE), building information modeling (BIM), and new learning environments to support AI curriculum.
- > **Research AI infrastructure:** This includes software and hardware that support the creation, tuning, and deployment of AI applications by research departments. Once the AI applications are developed, they can be run on provisioned infrastructure by grant stakeholders across the public and private sectors. These systems also support fundamental research related to new AI models and algorithms.
- > **Administrative and student services applications:** This includes software and hardware that support a smart campus, cybersecurity, and AI-enabled chatbots and student services that are increasingly adopted across university functions.

KEY CHALLENGES TO SUPPORTING AI IN HIGHER EDUCATION

- > **Growing resource demands:** Existing compute-specific architectures cannot support the growing demand for AI resources from nontechnical users.
- > **Complexity:** To support new workflows and workloads, infrastructure needs to be easy to manage, scalable, secure, and remote-ready.
- > **Siloed research labs, classrooms, and shadow AI:** Increasing operational overhead with new AI applications make IT management difficult, especially when additional infrastructure isn't compliant with IT standards and practices.

BENEFITS FOR HIGHER EDUCATION ORGANIZATIONS

- > **Optimized for AI development and deployment** with included AI frameworks and containers that save time and money, enabling faculty to focus on instruction and researchers to focus on scientific discovery
- > **Standardized and scalable infrastructure** deploys easily on popular data center platforms from VMware and Red Hat, mainstream NVIDIA-Certified Systems, and in the public cloud
- > **Supported by NVIDIA AI experts** to give IT organizations peace of mind when issues arise with included priority notifications, long-term support, and customized support-upgrade options

Higher education and research organizations need infrastructure that can enable not only their current applications but also future AI use cases. They need a foundational platform with working frameworks that simplifies, nurtures, and expands this new discipline.

AI-Ready Platform from NVIDIA

The NVIDIA AI Enterprise software suite enables universities and research organizations to harness the power of AI, even if they don't have AI expertise today. Optimized to streamline AI development and deployment, NVIDIA AI Enterprise includes proven, open-source containers and frameworks, certified to run on common data center platforms from VMware and Red Hat, mainstream NVIDIA-Certified Systems™ configured with GPUs or CPU-only, and the public cloud. Since support is included, higher education IT organizations benefit from the transparency of open source and the assurance that they can get help from the global NVIDIA Enterprise Support team whenever they need it. With the NVIDIA AI Enterprise software suite, AI is accessible at any scale, whether in a classroom, in a department, or across the entire university. The software suite provides the compute power, tools, and support necessary for institutions to complete their digital transformation journey.

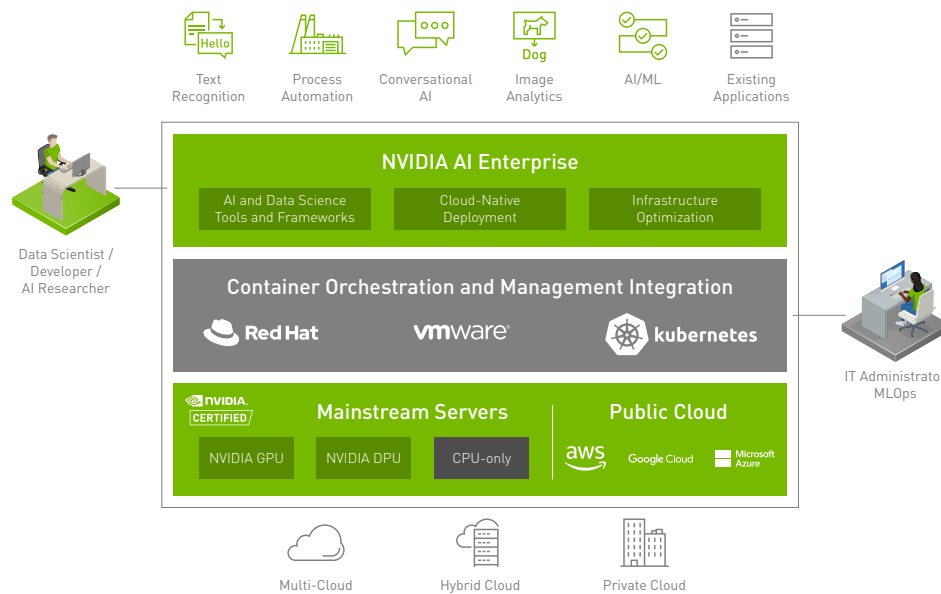


Figure 1. With the NVIDIA AI-Ready Platform, end users can access the software they need to build successful AI projects, and IT admins can support the projects using familiar tools.

Solution Highlights

Consolidate AI Applications for Higher Education

NVIDIA powers the traditional productivity, design, and visualization applications used throughout higher education organizations, in addition to AI applications and frameworks, which include:

- > Design, visualization, and business applications such as CAD, CAE, geospatial information systems (GIS), visual effects, and productivity solutions
- > Data science and research, including the NVIDIA DeepStream toolkit for video analytics, the NVIDIA® Riva SDK for conversational AI, and NVIDIA Triton™ Inference Server for the deployment of AI models at scale
- > Emerging AI applications to speed data analytics and automation and improve services, enabling a more modern, connected campus

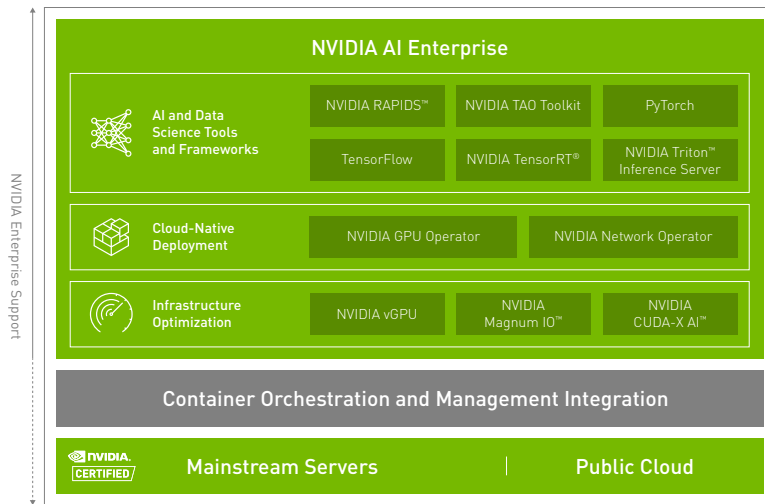


Figure 2. The NVIDIA AI Enterprise software suite includes the applications, frameworks, and tools used by AI researchers, data scientists, and developers, as well as tools for cloud-native deployments and infrastructure optimizations.

Create and Deliver AI with NVIDIA AI Enterprise

NVIDIA AI Enterprise is an end-to-end, cloud-native suite of AI and data analytics software, optimized so every higher education and research organization can succeed with AI. It’s also certified to deploy anywhere—from the enterprise data center to the public cloud—and includes global enterprise support and training.

NVIDIA AI frameworks and containers enable performance-optimized data science, training, and inference. They greatly simplify the building, sharing, and deployment of AI applications by eliminating the need to create and optimize specialized containers. Faculty can focus on instruction and researchers can achieve scientific breakthroughs more quickly. For example, NVIDIA RAPIDS makes it possible to streamline the data science process with up to 70X faster performance, while improving cost-effectiveness by up to 20X. Additionally, the NVIDIA TAO Toolkit boosts AI development by up to 10X, the NVIDIA Triton Inference Server allows organizations to improve TCO by 97X, and the NVIDIA A100 Tensor Core GPU enables them to achieve 100X lower latency when compared to CPU.

Deploy Anywhere

NVIDIA AI Enterprise can be deployed anywhere—including in a virtualized enterprise data center with VMware vSphere with Tanzu, on mainstream NVIDIA-Certified Servers, with the Red Hat OpenShift container platform, in the public cloud, and even on CPU-only servers. Tested and certified by NVIDIA with broadly adopted container, management, and server platforms, higher education organizations have the flexibility to deploy NVIDIA AI Enterprise with confidence in just about any data center environment—on premises, in a colocation facility, or in a private cloud. Organizations with a hybrid cloud strategy also have the flexibility to run NVIDIA AI Enterprise on GPU-accelerated public cloud instances on AWS, Azure, and Google with full NVIDIA support.

NVIDIA-Certified Systems

The NVIDIA AI Enterprise software suite is certified to run on NVIDIA-Certified Systems, which include the following:

- > **NVIDIA Ampere architecture-based GPUs** feature Tensor Core technology that delivers dramatic speedups to AI operations, reduces training times from weeks to hours, and provides massive inference acceleration.



Figure 3. NVIDIA-Certified Systems bring powerful speedups to AI training and inference.

NVIDIA-CERTIFIED SYSTEMS

- > Confidently deploy scalable hardware and software solutions that securely and optimally run accelerated workloads.
- > Learn more about accelerated servers at nvidia.com/certified-systems



- > **NVIDIA ConnectX® smart network interface cards (SmartNICs) and the NVIDIA BlueField® data processing unit (DPU)** provide a host of software-defined hardware engines for accelerating networking and security. These enable the best of both worlds: best-in-class AI training and inference performance with all the necessary levels of enterprise data privacy, integrity, and reliability.
- > **NVIDIA converged accelerators** combine the powerful performance of the NVIDIA Ampere architecture with the enhanced security and latency-reduction capabilities of the NVIDIA BlueField-2 DPU. Enterprises can use converged accelerators to create faster, more efficient, and secure AI systems in data centers and at the edge.

NVIDIA Enterprise Support

With NVIDIA AI Enterprise, higher education and research organizations get the transparency of open source with the assurance of full enterprise support and platform certification. They can extend their team to include NVIDIA experts, get support ticket prioritization and coordinated support across the full solution and partner products until resolution, control upgrades and maintenance schedules with long term support (LTS) options, and access the latest customer training and knowledge base resources.

NVIDIA AI Enterprise support includes:

- > Full enterprise-grade assistance for every deployment option, including bare metal, virtualized, containerized, GPU and CPU, and public cloud
- > Access to NVIDIA AI experts during local business hours for guidance on configuration and performance, including access to engineering
- > Priority notifications related to the latest security fixes and maintenance releases
- > Long-term support for up to three years for designated software branches
- > Customized support upgrade options, including a designated technical account manager (TAM) and Business Critical Support for 24x7 live agent access

NVIDIA AI Enterprise Trial Programs

NVIDIA offers the following trial programs that help customers evaluate products for free, depending on their existing infrastructure.

- > **NVIDIA LaunchPad**
NVIDIA LaunchPad is a program that provides organizations around the world with immediate, short-term access to the NVIDIA AI Enterprise software suite running on private accelerated computing infrastructure, including a set of curated hands-on labs for AI practitioners and IT staff. With NVIDIA LaunchPad, institutions can accelerate the development and deployment of modern, data-driven applications and quickly test and prototype the entire AI workflow on the same complete stack for purchase and deployment.
- > **Evaluation Software**
The NVIDIA AI Enterprise evaluation software is available for customers who have existing NVIDIA-Certified Systems. It's ideal for starting a proof-of-concept (POC) project for deployment at scale.

Build the Next Generation of Education with the AI-Ready Platform

The AI-Enterprise Ready Platform from NVIDIA offers end-to-end hardware and software that educational institutions need to digitally transform into an AI-driven university. Students, faculty, staff, and researchers can access the resources they need to build and use AI applications, and IT administrators can confidently and efficiently deploy applications and deliver uncompromised support to local and remote campuses using tools and infrastructure they know. This comprehensive solution is a key enabler of AI and the digital transformation occurring in higher education.

Ready to Get Started?

To learn more about NVIDIA AI for Enterprise and NVIDIA-Certified Systems, visit:

nvidia.com/ai-enterprise-suite

nvidia.com/certified-systems