SOLUTION OVERVIEW

# AI-READY ENTERPRISE PLATFORM FOR THE PUBLIC SECTOR

## Unleashing AI in the Public Sector with an End-to-End Platform

Artificial intelligence (AI) is transforming public services, from automating post mail sorting, to acting as a first-line response for citizen inquiries, to re-designing federal and civilian infrastructure. While the evolving needs of the public require that the varied functions of government invest in and explore AI that provides new services to meet them, limited budgets, lack of familiarity with data science tools, and siloed organizational infrastructures lead to struggles with identifying the strategy and platform that can enable success and deliver a rapid return on investment.

Outside of AI, the public sector continues to face pervasive, expansive security threats that increase the burden on IT teams. As more functions across services require remote access to applications, and new workloads are introduced, infrastructure silos often create disparate IT systems at risk for data leaks that pose national security risks and limit scaling. The sophisticated 2020 hack into the US federal treasury and commerce departments was only the largest and most visible example of a digital security breach, with hundreds of others expected to go unreported to the general public.

Security risks exacerbate the difficulties that companies face when it comes to incorporating AI, because unlike traditional enterprise applications, AI applications are a relatively recent development that challenge many IT departments. They're anchored in rapidly-evolving, open-source code and lack proven approaches that meet the rigors of scaled production settings and don't adhere to security management best practices. These difficulties mean that **only 53 percent of AI projects make it from pilot to production**. The complex integration of AI solutions with existing infrastructure is among the top three barriers to implementation.[1]

A government data center is built to provide resilient, always-up services to all the various roles in their department. There are three sets of systems that will challenge the public sector:

- Existing core applications: Computer aided design (CAD) and computer aided engineering (CAE) for building equipment, air- and water-craft vehicles, geospatial information systems (GIS) for mapping, training, and simulation, and building information modeling (BIM) for infrastructure development and more.
- New AI applications: Analytics and AI software that support the processing and management of citizen data for the IRS, postal, and veteran services, as well as new types of federal and civilian applications to enhance decisions and design better tools. Also included are radar processing and object detection, as well as natural language processing applications to engage with citizens as a first-line touchpoint.
- Emerging AI infrastructure: Software and hardware that support the creation, tuning, and deployment of AI applications by government departments.

### KEY CHALLENGES FOR AI WORKLOADS IN THE PUBLIC SECTOR

- Risk – Pulling together an end-to-end AI solution from disparate products—and integrating them with existing infrastructure—is difficult.
- Performance – Fast time-to-deployment and high performance are critical for AI, ML, and data analytics workloads.
- Scaling – Going from proof-of-concept to enterprise deployment requires effective scaling—through efficient use of resources—to ensure manageability, availability, and infrastructure cost management.

### BENEFITS FOR PUBLIC SECTOR ORGANIZATIONS

- Optimized to simplify AI development and deployment with included AI frameworks and containers, enabling insights to be gathered faster and business value to be delivered sooner
- Certified to deploy anywhere, including on popular data center platforms from VMware and Red Hat, mainstream NVIDIA-Certified Systems, and in the public cloud
- Supported by NVIDIA AI experts so AI projects stay on track, with included priority notifications, long-term support, and customized support upgrade options

1 "Gartner Identifies the Top Strategic Technology Trends for 2021," Gartner, October 2020

The public sector needs infrastructure that can encompass not only today's critical applications but also be AI-ready in the future. Organizations require a foundational platform that supports society's core services.

## AI-Ready Enterprise Platform from NVIDIA

The NVIDIA AI Enterprise software suite enables government agencies 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 on the public cloud. Since support is included, organizations benefit from the transparency of open source and the assurance that they can get help from the global NVIDIA Enterprise Support team to keep their AI projects on track.

With the NVIDIA AI Enterprise software suite, AI is accessible to organizations of any size, and provides the compute power, tools, and support that organizations need to focus on creating AI business value.

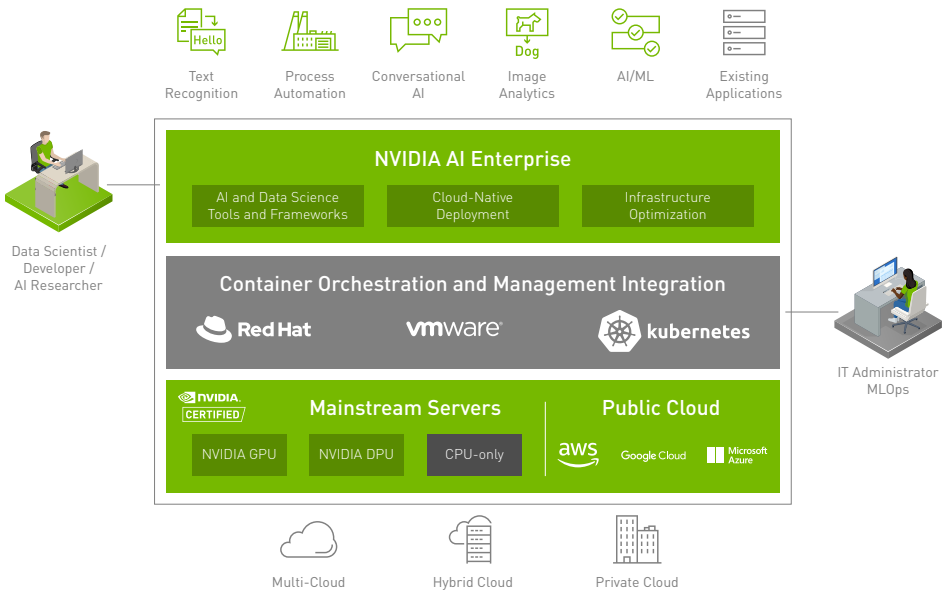


Figure 1. With the AI-Ready Platform, end users can access the software they need to build successful AI projects, and IT admins can support the projects using familiar tools.

## Solution Highlights

### Consolidate AI Applications for the Public Sector

NVIDIA powers the traditional office productivity, and 3D design and visualization applications required by public sector organizations, in addition to AI applications and frameworks, some examples of which include:

- Design, visualization, and business applications, including advanced visualization applications like CAD, CAE, and GIS, 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 machine learning to improve the services that enable our society to thrive.

## NVIDIA AI Enterprise Software Suite

NVIDIA AI Enterprise is an end-to-end, cloud-native suite of AI and data analytics software, optimized so every government agency 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 also simplify the building, sharing, and deployment of AI applications so government agencies can gather insights faster and deliver business value sooner. For example, NVIDIA Rapids™ makes it possible for organizations to streamline the data science process with up to 70X faster performance, while improving cost-effectiveness by up to 20X. Additionally, NVIDIA TAO Toolkit boosts AI development by up to 10X, the NVIDIA Triton Inference Server allows organizations to improve TCO by 97X, and NVIDIA A100 enables them to achieve 100X lower latency when compared to CPU.

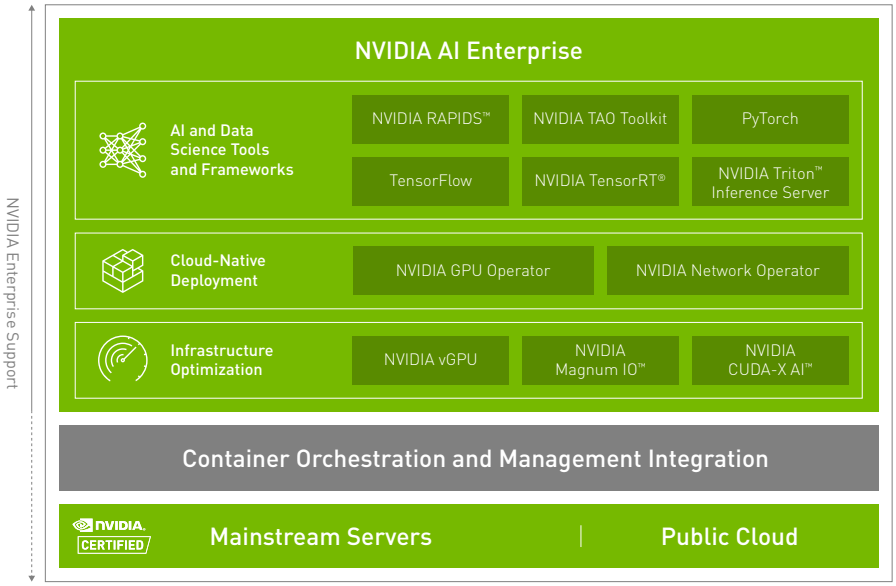


Figure 2. The NVIDIA AI Enterprise software suite includes the applications, frameworks, and tools used by AI researchers, data scientists, and developers, as well as tools for cloud-native deployments and infrastructure optimizations.

## NVIDIA-Certified Systems

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

- **NVIDIA Ampere architecture-based GPUs. The Tensor Core technology included in the Ampere architecture delivers dramatic speedups to AI operations, reduced training times from weeks to hours, and provides massive inference acceleration.**
- **NVIDIA® ConnectX® SmartNICs and the NVIDIA BlueField® data processing unit (DPU). These provide a host of software-defined hardware engines for accelerating networking and security. They also 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: These 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.

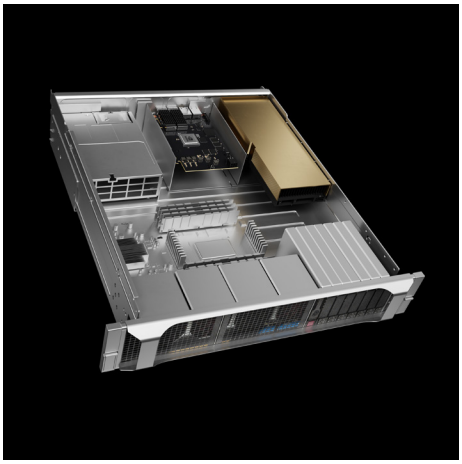

Figure 3. NVIDIA-Certified Systems bring powerful speedups to AI training and inference.

### 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 CERTIFIED

**NVIDIA Enterprise Support**

With NVIDIA AI Enterprise, government agencies 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 3-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 enable customers to evaluate products for free, based on their existing infrastructure:

- NVIDIA LaunchPad
  **The NVIDIA LaunchPad program provides worldwide enterprises and organizations with immediate, short-term access to the NVIDIA AI Enterprise software suite running on private accelerated computing infrastructure. It also includes a set of hands-on labs for AI practitioners and IT staff. IT administrators learn best practices for deploying NVIDIA AI Enterprise, and AI practitioners learn how to optimize training and inferencing workloads using AI tools and frameworks.**
- Evaluation Software
  **NVIDIA AI Enterprise evaluation software is available for customers with existing NVIDIA-Certified Systems who are ready to start a proof-of-concept (POC) project for deployment at scale.**

## Build the Next Generation of Government with the AI-Ready Platform

The AI-Enterprise Ready Platform from NVIDIA offers end-to-end hardware and software that government agencies need to modernize and accelerate their infrastructure. Developers, data scientists, and researchers can access the resources they need to build and deploy AI applications efficiently, and IT administrators can confidently deliver uncompromised support with tools and infrastructure they know well. This comprehensive solution pushes government agencies forward to serve the public and improve mission-critical outcomes.

### 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**



Deloitte. NVIDIA.