

# Get off the whiteboard and into production on real GPUs, fast.

## “AI Factory” Accelerator



Many AI initiatives stall without a clear use case and a scalable software-to-compute foundation. Organizations know they must operationalize AI, but they often struggle to select the right starting point, validate performance on their infrastructure and prove value quickly. CDW’s “AI Factory” Accelerator removes uncertainty by identifying a prioritized “hero” use case and delivering a time-boxed validation of an NVIDIA Inference Microservices Blueprint on NVIDIA DGX. Using a small dataset and light customization, we deploy the blueprint, measure real performance and deliver a production-ready proof of concept (POC) with a clear roadmap — reducing risk, accelerating adoption and driving measurable business impact.

CDW’s “AI Factory” Accelerator can help you achieve:



**Agility**



**Faster Innovation**



**Reduced Risk**

### Finding the Right Solution

CDW delivers AI solutions through proven architecture, secure integration and optimized performance, backed by deep technical expertise and strategic partnerships with leading technology providers. This service establishes a strong foundation for scalable AI deployment and delivers a clear, business-aligned roadmap. By identifying high-impact business objectives suited for AI, organizations gain greater control over AI initiatives and reduce uncertainty. A cohesive, time-boxed proof of concept validates technical feasibility, demonstrates measurable performance and builds confidence that AI projects can deliver rapid ROI.

The service includes: Gap analysis; use-case identification; NVIDIA AI Blueprint POC deployment on DGX; micro-dataset ingestion (1-2 GB or up to ~5k documents); minimal configuration tuning and basic guardrails; metrics capture and rapid evaluation; demo scripts; runbook; and an executive readout with a staged roadmap.

The outcome of this engagement includes a prioritized “hero” use case, infrastructure readiness assessment and validated NVIDIA Inference Microservices Blueprint on NVIDIA DGX.

## CDW + NVIDIA

CDW and NVIDIA have a strong partnership focused on delivering cutting-edge AI, visualization and high-performance computing solutions to businesses. As an NVIDIA Elite Solution Provider, CDW combines deep industry expertise with NVIDIA’s advanced technologies — such as GPU-powered AI platforms, DGX systems and AI Enterprise software — to help organizations accelerate innovation and unlock the full potential of their data. In 2025, CDW achieved the NVIDIA DGX SuperPOD Partner Designation for scalable AI infrastructure delivery.

Working with CDW provides access to tailored infrastructure design, secure deployment strategies and end-to-end support for scaling AI initiatives. This collaboration enables businesses to achieve faster ROI, streamline operations and build sustainable, production-ready AI environments that drive competitive advantage.

CDW’s full lifecycle of Services can support your organization no matter where you are on your journey



## Services Overview

The following stages are included in CDW's "AI Factory" Accelerator engagement:

Engagement Stage	Stage Details
Kickoff and Guardrails (Use Case + Infrastructure Readiness Workshop)	<ul style="list-style-type: none"><li>• Ideate and prioritize feasible AI use cases and subsequently select a "hero" use case.</li><li>• Select the appropriate NVIDIA Inference Microservices Blueprint (for next phase).</li><li>• Define micro-dataset and redactions.</li><li>• Set access and RBAC.</li></ul>
Baseline NVIDIA Blueprint Deployment	<ul style="list-style-type: none"><li>• Installation and configuration on NVIDIA DGX systems (or optionally on approved select OEM AI pre-engineered systems with NVIDIA GPUs).</li><li>• Use of customer's existing NVIDIA AI Enterprise subscription.</li><li>• Light customization to match the selected use case and AI infrastructure configuration and capacity.</li></ul>
Micro-Dataset Ingestion (1-2 GB or up to ~5k docs)	<ul style="list-style-type: none"><li>• Use of a small, representative customer dataset.</li><li>• Minimal preprocessing sufficient for blueprint validation.</li></ul>
Polished Proof of Concept (POC)	<ul style="list-style-type: none"><li>• "Ready for First Use" POC environment.</li><li>• Minimal configuration tuning and basic guardrails.</li><li>• Golden-path scripts, light branding and role-based demo accounts.</li></ul>
Performance Measurement and Analysis	<ul style="list-style-type: none"><li>• Capture utilization, latency and throughput.</li><li>• Quick retrieval eval (10-50 queries).</li><li>• Optional baseline comparison.</li></ul>
Final Deliverables Package	<ul style="list-style-type: none"><li>• POC scripts and runbook for blueprint redeployment.</li><li>• Metrics report.</li><li>• Executive readout with roadmap to production:<ul style="list-style-type: none"><li>• Observations on scalability and AI GPU infrastructure environment readiness.</li><li>• Staged steps for scaling to a full production implementation including risks and prerequisites.</li></ul></li></ul>

**To learn more about CDW's "AI Factory" Accelerator, contact your account manager or call 800.800.4239.**

