

Your Way to AI

Dell AI Factory with NVIDIA: Reimagine the speed of business with cutting-edge AI innovation and scalable AI workflow automation



The Dell AI Factory with NVIDIA

Accelerating AI adoption allows you to extract the full potential of your data

The Dell AI Factory with NVIDIA speeds AI adoption by delivering integrated Dell and NVIDIA® capabilities to accelerate your AI-powered use cases, integrate your data and workflows and enable you to design your own AI journey for repeatable, scalable outcomes.



Faster time to value



Transform insights



Optimize productivity



Deliver trusted outcomes

The world's broadest GenAI solutions portfolio from desktop to data center to cloud, all in one place.²

75%

more cost-effective inferencing than public cloud IaaS.³

86%

faster time to value via turnkey, full-stack solutions with services.⁴

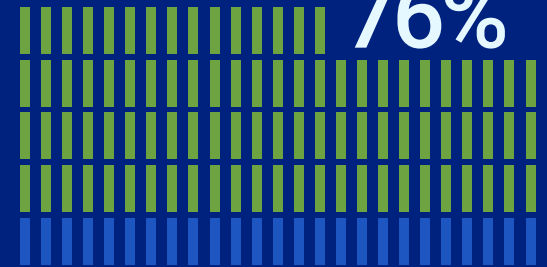
¹ Dell Technologies Innovation Catalyst Study, February 2024.

² Based on Dell analysis, August 2023. Dell Technologies offers solutions engineered to support AI workloads from workstations PCs (mobile and fixed) to servers for high-performance computing, data storage, cloud-native software-defined infrastructure, networking switches, data protection, HCI and services.

³ Based on Enterprise Strategy Group research commissioned by Dell, comparing on-premises Dell infrastructure versus native public cloud infrastructure as a service, April 2024. Analyzed models show a 7B parameter LLM leveraging RAG for an organization of 5k users being up to 38% more cost effective while a 70B parameter LLM leveraging RAG for an organization of 50k users being up to 75% more cost effective. Actual results may vary.

⁴ Estimate based on Dell analysis in May 2024 comparing time to set up a 2-node Kubernetes cluster for a general-purpose LLM using automated scripts vs deploying a common design manually. Setup time includes base installation only. Actual setup time will vary depending on solution configuration.

76%



of IT and business leaders believe GenAI will deliver transformative value for their organization.¹

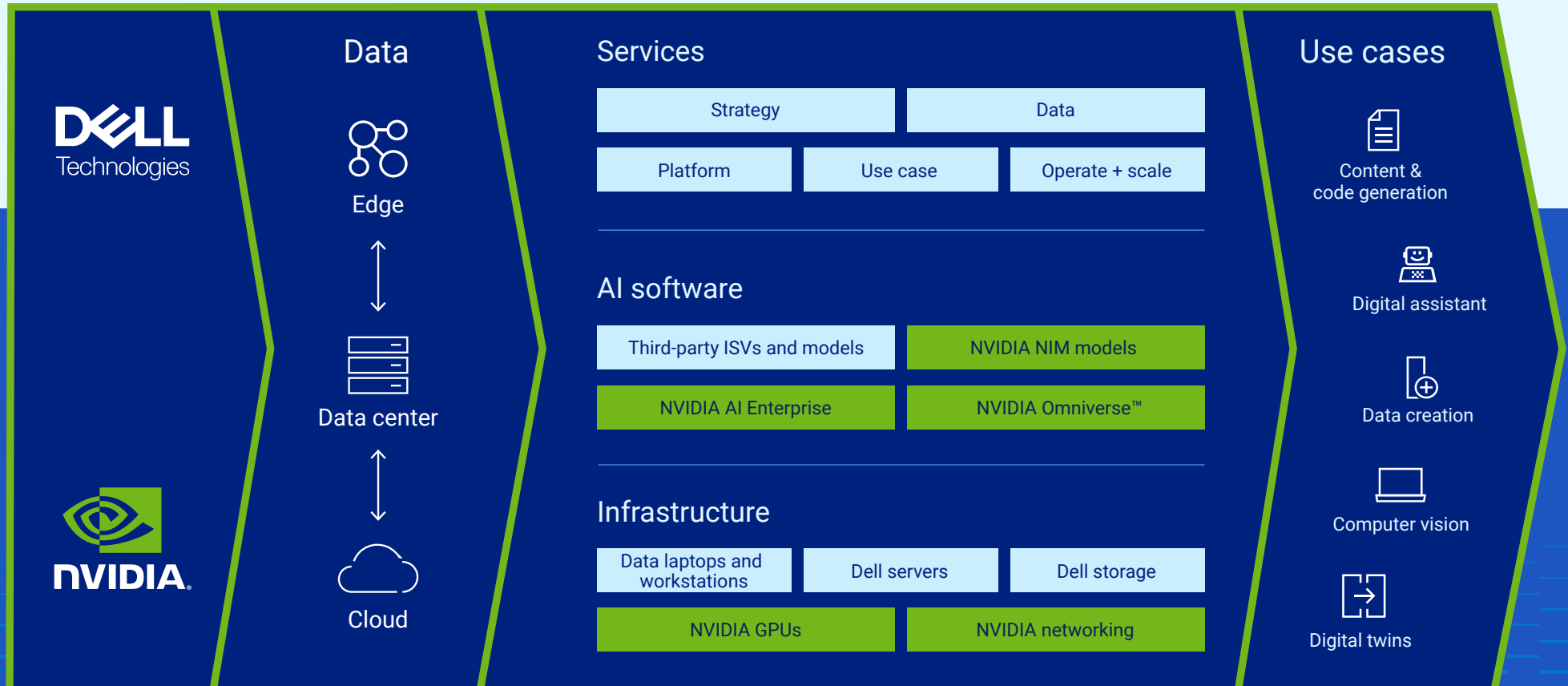
Dell Technologies and NVIDIA have a long-standing partnership with **more than 25 years of joint innovation**, focused on accelerating innovation and delivering cutting-edge platforms, solutions and software that enable transformative results for our joint customers.

The Dell AI Factory with NVIDIA

The industry's first end-to-end enterprise AI solution

The Dell AI Factory with NVIDIA delivers a comprehensive portfolio of AI technologies, validated and turnkey solutions, with expert services to help you achieve AI outcomes faster.

Extend your enterprise with AI and GenAI at scale, powered by the broad Dell portfolio of AI infrastructure and services with NVIDIA industry-leading accelerated computing, a full stack that includes GPUs; networking; and NVIDIA AI Enterprise software, NVIDIA Inference Microservices (NIM), models and agent blueprints.

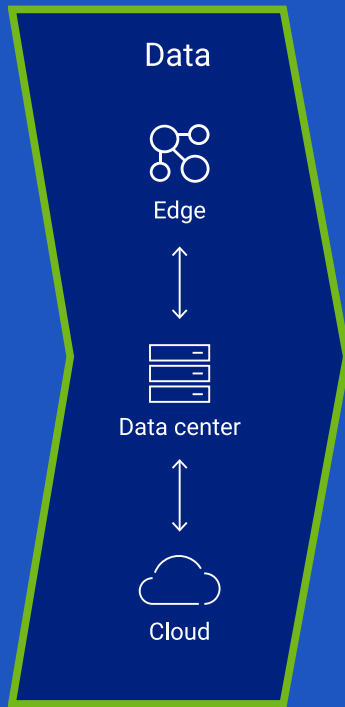


Your data is your differentiator

AI starts with better data management and preparation. The ability to accelerate business decisions through the utilization of robust, accurate and comprehensive data sets will distinguish industry leaders from the rest.

Data fuels the AI factory

Your most valuable data often resides on-premises and at the edge – places where you have exclusive ownership. This can minimize latency and ensure sensitive information remains secure.



Bring AI to your most valuable data with leaders in storing, protecting and managing that data

Your data management journey is not a one-time task but an ongoing process, utilizing a set of practices and tools to help your organization continuously unlock value from your data. Effective data management is crucial for any successful generative AI initiative.

Dell Technologies, in collaboration with NVIDIA, helps organizations like yours navigate the data management journey with ease. In the realm of AI-driven data workloads, the journey is just as important as the destination.

The Dell AI Factory with NVIDIA works with Dell data lakehouse strategies to help manage and visualize data and be prepared for GenAI ingestion with Dell Services and partners. Data storage is enhanced further with AI-ready Dell PowerScale, delivering high-performance, certified platforms for AI Factory as well as NVIDIA DGX Pod, DGX Superpod and OVX systems.

Outcomes powered by use cases

Simplify the deployment of your most important AI uses cases with validated solutions and tailored services.



Content and code generation

- Curate branded content.
- Enhance content personalization.
- Update market reports.
- Aggregate industry news.
- Monitor competitive activities.
- Integrate enterprise systems.
- Automate testing processes.
- Facilitate code refactoring.
- Enhance security protocols.
- Streamline database queries.



Digital assistant

- Provide customer-facing support.
- Utilize training and education.
- Optimize scheduling with an assistant.
- Track project deadlines.
- Automate routine tasks.



Data creation

- Enhance business intelligence.
- Streamline project research.
- Facilitate legal queries.
- Optimize procurement searches.
- Improve financial analytics.



Computer vision

- Detect theft.
- Visually inspect equipment and products.
- Automate mechanical operations.
- Analyze worker safety.
- Monitor road conditions.



Digital twins

- Customer experiences.
- Product development and prototyping.
- Maintenance and operations.
- Logistics and supply chain.
- Training and guidance.

Use cases



Content & code generation



Digital assistant



Data creation



Computer vision



Digital twins

AI-optimized infrastructure

Infrastructure is the foundation of the AI factory. Right-size your AI investment with the flexibility to run your workloads anywhere through the broad Dell AI portfolio from desktop to data center.

Identify the use case

What are the infrastructure requirements for your desired business outcome?



Decide what AI model/type you need

With an understanding of your business goals and use case, you can determine the right balance of performance and cost.



Right-size your investment

Some models can be deployed on laptops, some will require a large number of GPUs, like the 8x GPUs featured on the Dell PowerEdge XE9680. It really depends on what you're trying to accomplish. This is where the AI factory framework becomes helpful.

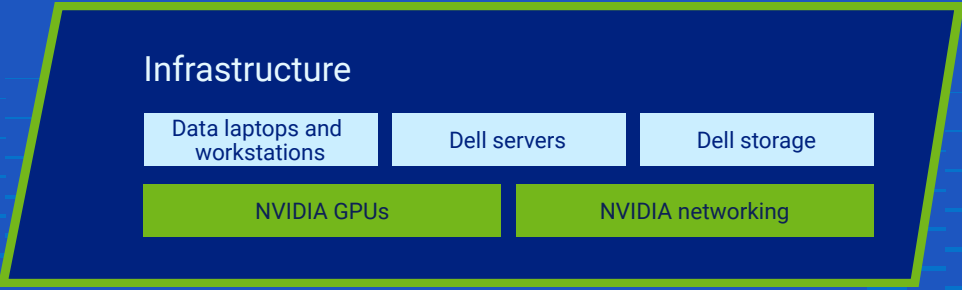
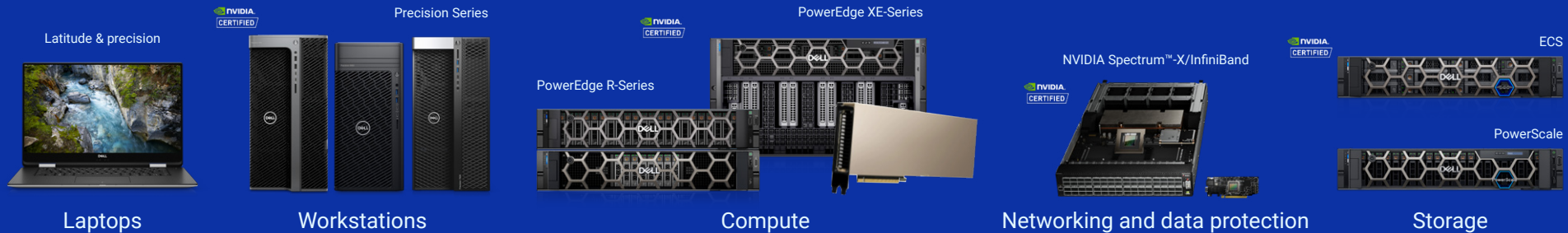
Product highlights

Dell NativeEdge is the first edge orchestration platform that automates the delivery of NVIDIA AI Enterprise at the edge.⁵

The **Dell PowerEdge XE9680L server** offers direct liquid cooling and 8 next-generation NVIDIA Blackwell Tensor Core GPUs in a compact 4U form factor.

Dell PowerScale is the world's first ethernet-based NVIDIA DGX SuperPOD certified storage solution.

Infrastructure



⁵ Based on Dell analysis, May 2024.

NVIDIA AI Enterprise – the Operating System for AI

Flexible and robust end-to-end platform for enterprise AI



AI software

Third-party ISVs and models

NVIDIA NIM models

NVIDIA AI Enterprise

NVIDIA Omniverse™

Fast-track your AI adoption – desktop to data center

Quickly get up and running and give your teams hands-on experience with a local prototype environment. Start building organization skills with a clear understanding of data pipelines and business processes, enabling you to seamlessly scale as your needs grow.



Simple development launchpad

Perform local prototyping in a safe and secure environment.



Develop and prepare to scale

Rapidly build AI and data workflows with container-based microservices.



Deploy in the enterprise

Orchestrate and optimize with a scalable infrastructure framework.



AI workstations

Precision



Compute

PowerEdge & NVIDIA GPUs



Storage

PowerScale/PowerStore

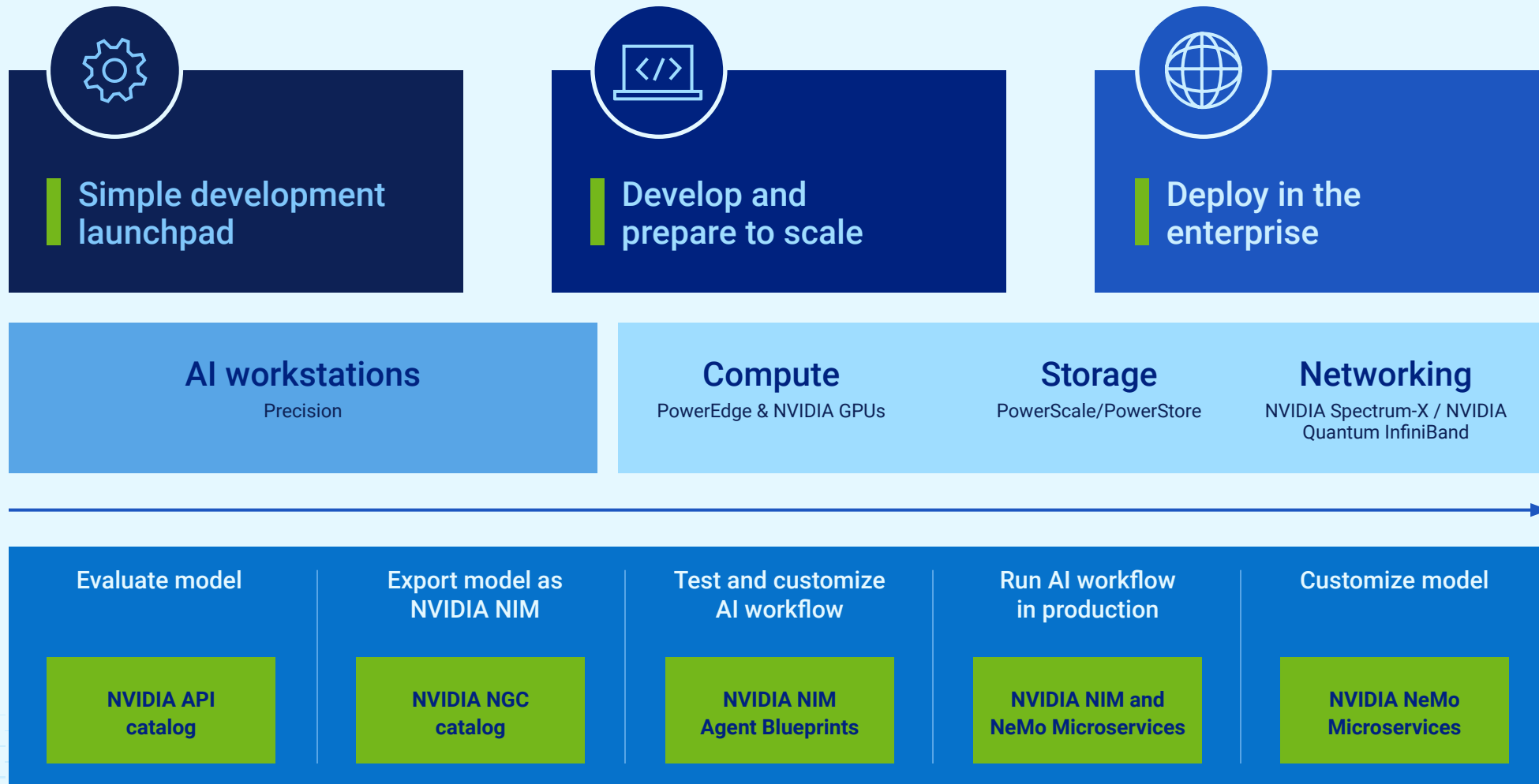


Networking

NVIDIA Spectrum-X /
NVIDIA Quantum InfiniBand

Scale as you grow with a broad portfolio designed for a wide range of workloads.

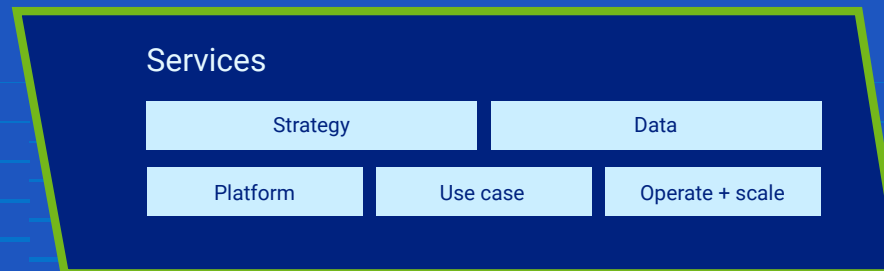
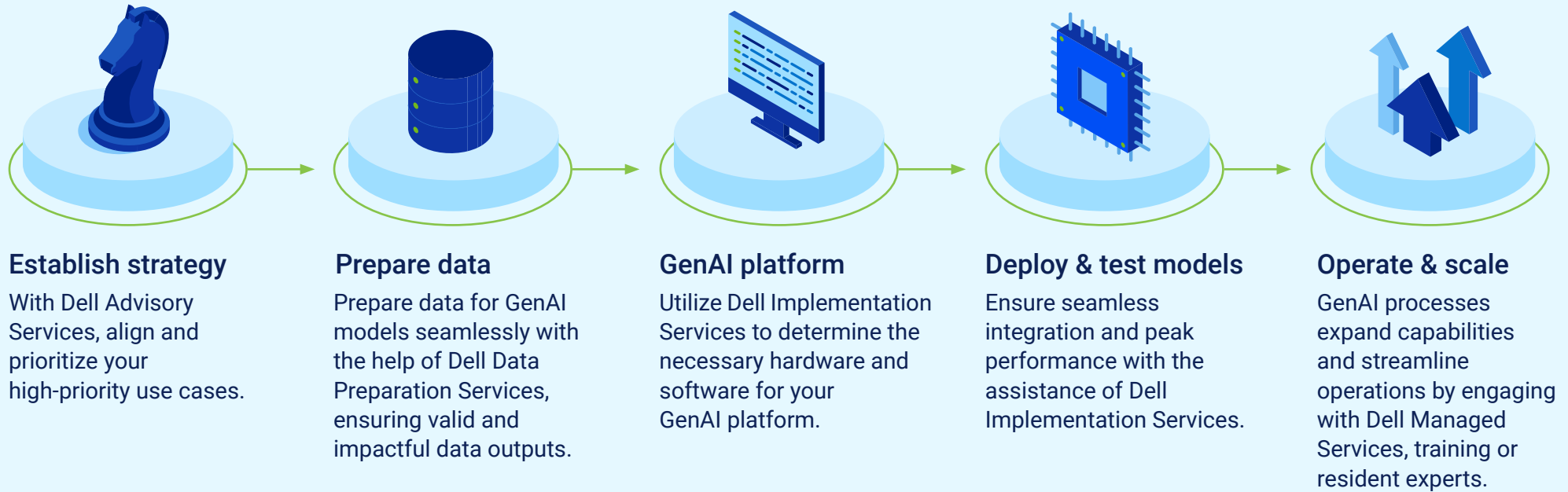
Combined, scalable infrastructure and microservices architecture



Take the guesswork out of building your own GenAI tools with open source. Instead, take advantage of having these different aspects integrated and connected, making it simpler and easier to find success with a comprehensive solution for deploying GenAI applications at scale.

Effectively embracing generative AI is a journey

Dell Technologies has extensive experience guiding customers through their AI journeys so you can quickly accelerate outcomes by aligning business objectives with the right technical solutions at scale.



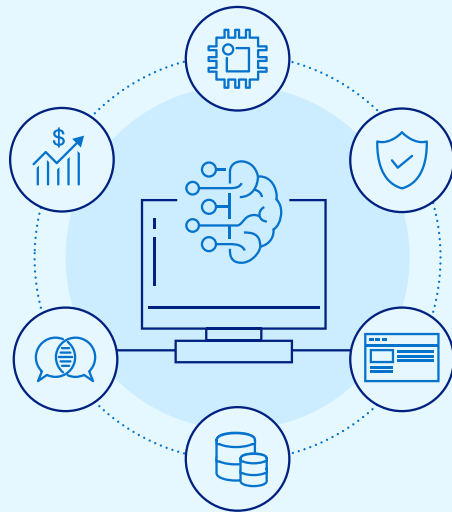
Two ways to get started today

Rapidly start delivering organizational value with Dell AI Factory with NVIDIA.

Strategic planning

Accelerator Workshop for Generative AI

- Start your journey by developing a winning strategy with a fee-waived half-day workshop.
- Address challenges and gaps, prioritize objectives and identify opportunities.
- A more complete readiness assessment is also available for a deeper dive into infrastructure requirements, AI models, operational integrations and more.



Technical preparation

Jump-start your journey to success with an easy-to-use mobile lab – including a Dell Mobile Precision Workstation and two days of Consulting Services to help get you started.

- Portable GenAI testing and demonstration.
- Rapid prototyping in a prevalidated environment.
- Cost-effective, low-risk exploration of GenAI use cases.
- Closing GenAI skills gaps.



Dell Mobile Precision Workstation 5690/7780 with NVIDIA GPUs



A sandbox environment for GenAI experimentation
Deployed and configured RAG architecture

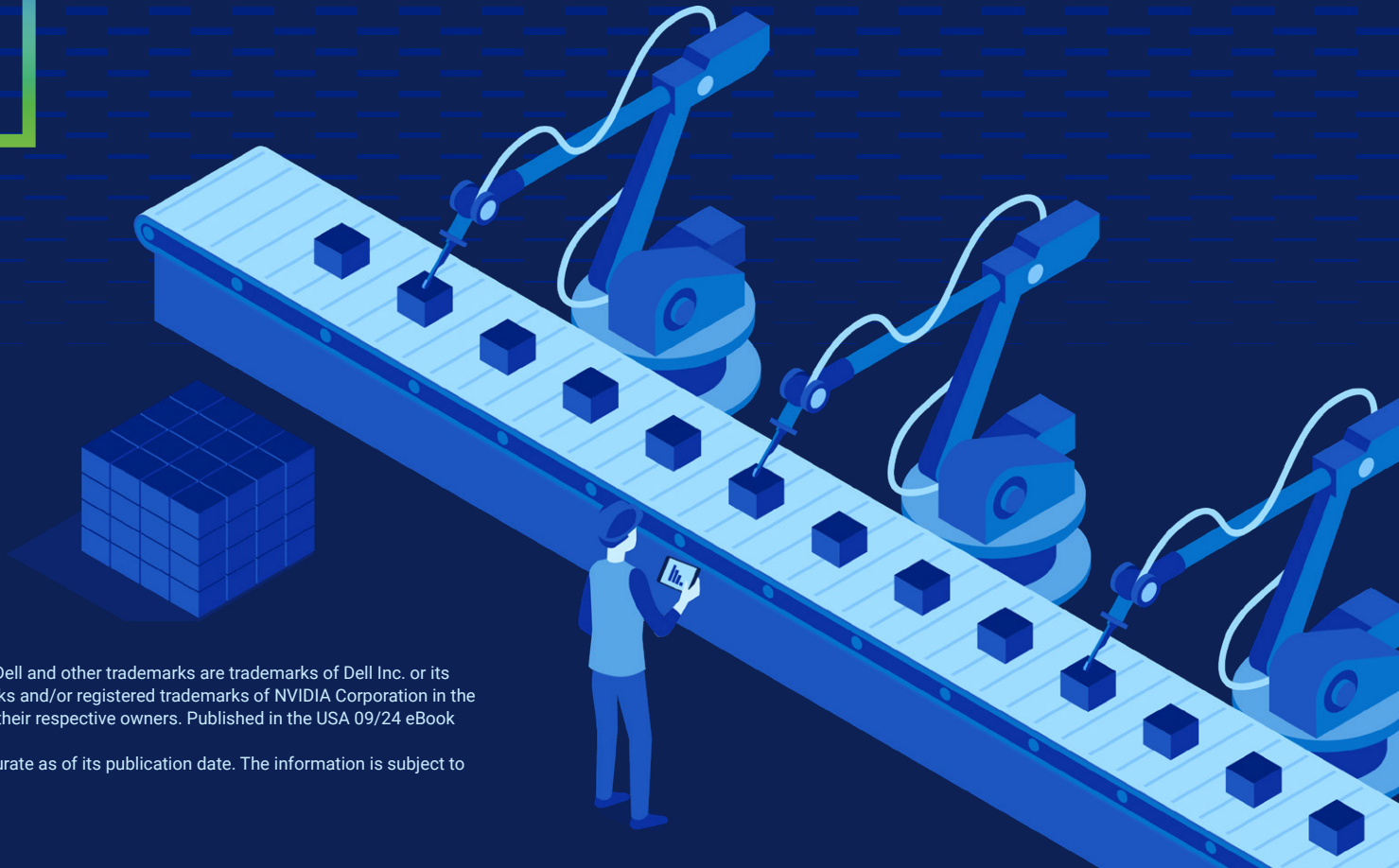
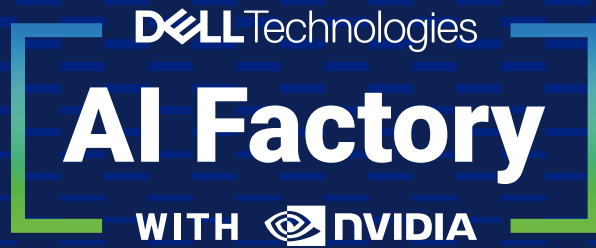
NVIDIA AI Workbench platform ready for developers
Preconfigured to run software and validated for performance

Foundational use case to enable further exploration
Initial chatbot use case implemented with your data

Partner with Dell Technologies and NVIDIA to unlock the full potential of GenAI

Visit our website Dell.com/en-us/dt/nvidia-ai.htm

to learn more and start your AI journey today.



Copyright © 2024 Dell Inc. or its subsidiaries. All Rights Reserved. Dell and other trademarks are trademarks of Dell Inc. or its subsidiaries. NVIDIA®, Omniverse™, and Spectrum™-X are trademarks and/or registered trademarks of NVIDIA Corporation in the U.S. and other countries. Other trademarks may be the property of their respective owners. Published in the USA 09/24 eBook

Dell Technologies believes the information in this document is accurate as of its publication date. The information is subject to change without notice.