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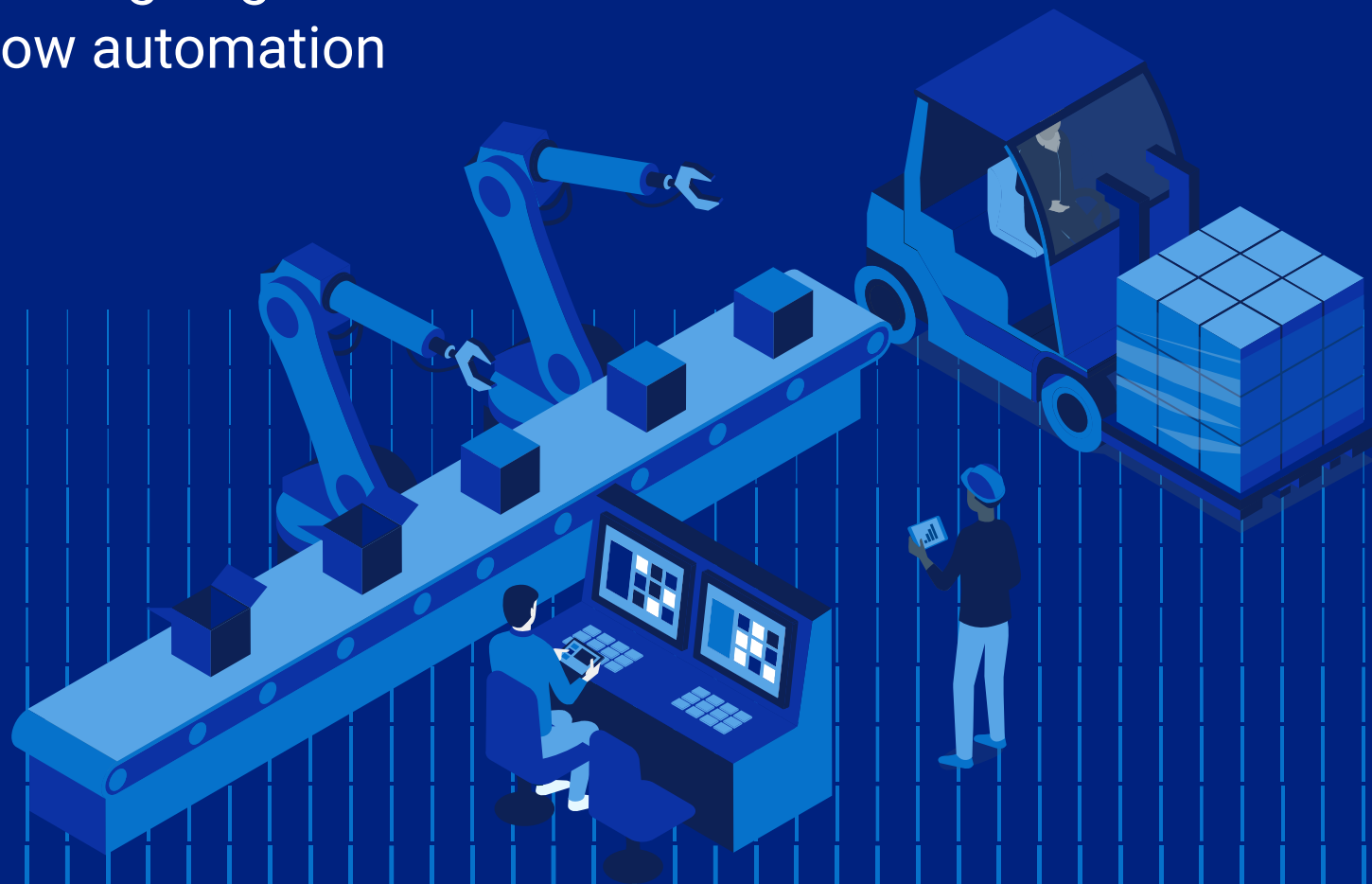
Dell AI Factory with NVIDIA: Reimagine the speed of business with cutting-edge AI innovation and scalable AI workflow automation

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.

Drive growth

Enhance experiences

Optimize productivity

Accelerate innovation

75%
of organizations say AI/GenAI has become key to their business strategy.[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.

93%
of organizations experience challenges when integrating AI/GenAI into their business strategies[1]

60%
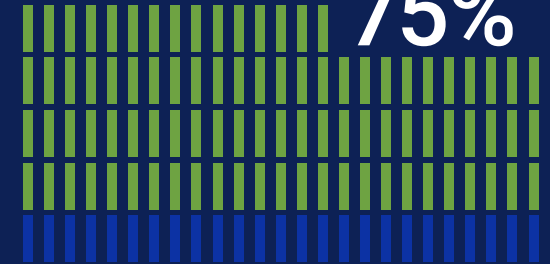savings with Dell AI Factory vs. public cloud[2]

29%
of organizations are already using agentic AI and another 44% are planning adoption within the next year[3]

The Dell AI Factory with NVIDIA

Accelerate the AI journey

The Dell AI Factory with NVIDIA is the industry's first end-to-end enterprise AI solution. It 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, GenAI and agentic AI 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 NIM and NeMo Microservices, models and agent blueprints.

DELL Technologies

AI
-
Generative AI
-
Agentic AI

NVIDIA

Data

Edge

Data center

Cloud

AI Services

Strategic guidance

Data management

Rack integration for GPUs

Use case implementation

Managed operations

AI software

Third-party ISVs and models

NVIDIA NIM and NVIDIA NeMo microservices

NVIDIA AI Enterprise and NVIDIA Run:AI

NVIDIA Omniverse™

Accelerated infrastructure

High-performance PCs

Servers

Storage and cyber resilience

Networking

GPUs and CPUs

Use cases

Content & code generation

Digital assistant

Data creation

Computer vision

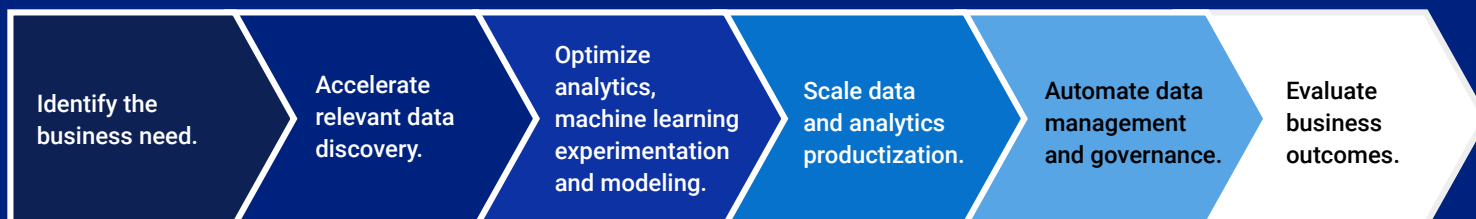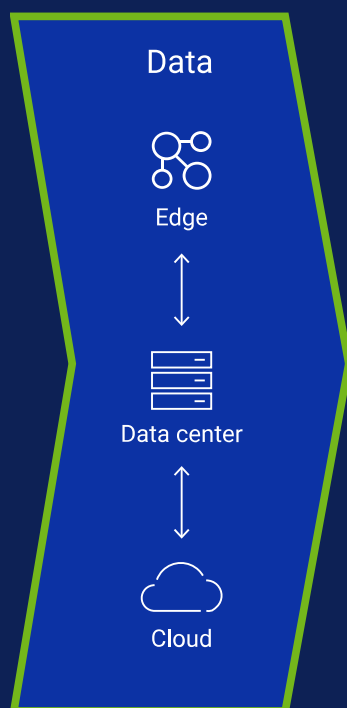Digital twins

Sustainability | Security

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.

Data

Edge

↕

Data center

↕

Cloud

Identify the business need.	Accelerate relevant data discovery.	Optimize analytics, machine learning experimentation and modeling.	Scale data and analytics productization.	Automate data management and governance.	Evaluate business outcomes.

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 agentic 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 agentic AI 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 to cloud.

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.

Infrastructure

Dell Pro Max
Jumpstart AI development on-prem with the #1 workstations in the world,[4] that feature NVIDIA's most powerful professional graphics

Dell PowerEdge servers
Maximize AI acceleration and future-proof your AI infrastructure with NVIDIA B300 and RTX Pro Blackwell Server GPUs

PowerScale & Dell AI Data Platform with NVIDIA
Seamless access to structured and unstructured data—with support for NVIDIA RAG, VILA, Llama Nemotron and built-in connectors and integrations

Dell PowerSwitch SN5600
High-performance AI fabric with 64x 800GbE delivers speed at scale—backed by a seamless, end-to-end customer experience

Accelerated infrastructure

High-performance PCs	Servers	Storage and cyber resilience
Networking		GPUs and CPUs

[4] Based on IDC Worldwide Quarterly Workstation Tracker, Feb 2025 – Units.

Kickstart AI development locally with Dell Pro Max

Exceptional performance
Engineered and configurable to power your AI tasks effortlessly

Certified reliability
100+ ISV applications tested for seamless performance

Unmatched security
The industry's most secure & manageable commercial AI PCs

Scalability for your AI goals
Functional infrastructure within the Dell AI Factory with NVIDIA

Dell Pro Max high-performance PCs empower AI development and innovation from desk to data center. Ideal for local AI prototyping, they reduce cloud costs and security concerns, while enabling you with the tools to build and experiment confidently with NVIDIA AI software platforms.

Start at the edge with Dell Pro Max PCs and scale effortlessly to enterprise-level AI through the Dell AI Factory with NVIDIA, streamlining integration and accelerating growth as your AI needs evolve.

Tailored AI power for every workflow
Dell Pro Max high-performance PCs with NVIDIA RTX Pro Blackwell graphics

Tiny powerhouse, massive AI power
Dell Pro Max with GB10 Powered by NVIDIA Grace Blackwell architecture

Data center-level AI power at the deskside
Dell Pro Max with GB300 Powered by NVIDIA Grace Blackwell architecture

NVIDIA AI Enterprise — the Operating System for AI

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

MLOps

AI applications

NVIDIA AI Enterprise

AI microservices

AI workflows

AI application frameworks

AI development and deployment tools

AI infrastructure management and optimization

Enterprise support

Security patching

Branch support

Stability

Accelerated Infrastructure

Workstation

Data center

Edge

Cloud

The cloud-native software platform is designed to accelerate data science pipelines and streamline the development and deployment of production-grade AI applications.

Easy to use microservices provide optimized model performance and industry-standard APIs to ensure a smooth transition from prototype to production.

Secure software lifecycle, end-to-end manageability and enterprise support enable organizations to confidently build and run AI solutions.

AI software

Third-party ISVs and models

NVIDIA NIM and NVIDIA NeMo microservices

NVIDIA AI Enterprise and NVIDIA Run:AI

NVIDIA Omniverse™

Fast-track your AI adoption — desktop to data center to cloud

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.

High-performance PCs & Software
Dell Pro Max & Dell Pro AI Studio

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

Simple development launchpad

Develop and prepare to scale

Deploy in the enterprise

High-performance PCs & Software
Dell Pro Max & Dell Pro AI Studio

Compute
PowerEdge & NVIDIA GPUs

Storage
PowerScale/PowerStore

Networking
NVIDIA Spectrum-X / NVIDIA Quantum InfiniBand

Evaluate model	Export model as NVIDIA NIM	Test and customize AI workflow	Run AI workflow in production	Customize model
NVIDIA API catalog	**NVIDIA NGC catalog**	**NVIDIA NIM Agent Blueprints**	**NVIDIA NIM and NeMo Microservices**	**NVIDIA NeMo Microservices**

Take the guesswork out of building your own AI 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 and agentic AI applications at scale.

The rise of agentic AI

By 2028, 33% of enterprise software applications will include agentic AI,[5] redefining how businesses work and compete. According to Forrester, agentic AI represents more than just an efficiency play — it signals a fundamental shift in how enterprises compete, operate, and deliver value.[6]

So what is agentic AI? Unlike traditional or even generative AI, agentic AI is a**utonomous, adaptive and goal oriented.** It doesn't just follow commands — it takes initiative, learns from context and optimizes its actions to meet business objectives. It's like an intelligent collaborator capable of **enhancing human decision-making,** automating repetitive tasks and driving continuous improvement across workstreams.

Agentic AI allows organizations to **embed business knowledge, data and procedures into intelligent systems** that act with purpose. As the workforce evolves, employees will increasingly coordinate a network of AI agents, each handling different tasks—from customer service to supply chain optimization.

Agentic AI

Coding assistant	Customer service	Patient Care

Cybersecurity	Supply chain management

Dell and NVIDIA provide the foundational tools to bring agentic AI to life:

NVIDIA Blueprints for a fast start

NVIDIA NeMo Framework for large language model customization

NVIDIA NIM microservices for scalable AI deployment

The generative AI and agentic AI 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.

Establish strategy

With Dell Advisory Services, align and prioritize your high-priority use cases.

Prepare data

Prepare data for agentic AI models seamlessly with the help of Dell Data Preparation Services, ensuring valid and impactful data outputs.

Agentic AI platform

Use Dell Implementation Services to determine the necessary hardware and software for your agentic AI platform.
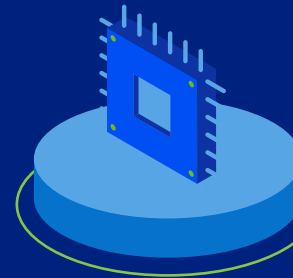
Deploy & test models

Ensure seamless integration and peak performance with the assistance of Dell Implementation Services.
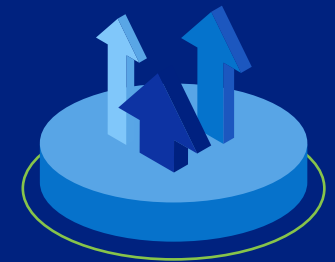
Operate & scale

Agentic AI processes expand capabilities and streamline operations by engaging with Dell Managed Services, training or resident experts.

AI Services

Strategic guidance	Data management

Rack integration for GPUs	Use case implementation	Managed operations

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

Dell Precision Mobile workstation 5690/7780 with NVIDIA GPUs

A sandbox environment for AI 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 and agentic AI

For more information,
visit our <u>The Dell AI Factory with NVIDIA | Dell USA</u>

DELLTechnologies

AI Factory

WITH ◁▷ NVIDIA

Additional resources:

- <u>Services Accelerator Workshop</u>
- <u>AI Solutions Explorer</u>
- Turbocharge Your Way to AI <u>infographic</u>
- Dell AI Factory with NVIDIA <u>video</u>
- Dell AI Factory with NVIDIA use case <u>video</u>
- ESG <u>white paper</u>