



Accelerating AI innovation  
with inferencing at the edge.





# Table of Contents

3

**Introduction**  
The Evolution of Artificial Intelligence  
Generative and Conversational AI: The New Frontier  
The Importance of Edge AI Inferencing

4

**Inferencing at the Edge:  
A Trillion-Dollar Opportunity**  
Key Benefits of Edge AI  
Real-Time Decision Making and Security  
Cost and Energy Efficiency at the Edge

6

**Industry-Specific Applications**  
Transformative Use Cases Across Industries  
Layering AI and GenAI for Holistic Solutions

7

**Customer Success Stories**  
Railcar Inspection: Duos Technologies  
Formula 1 AI Innovations: McLaren Racing  
AI-Powered Medical Diagnostics: Anke

8

**Simplifying Your Edge AI Journey**  
Overcoming Challenges with Dell Technologies  
The Dell AI Factory: Data, Infrastructure, Ecosystem,  
and Services  
Edge-Optimized AI Infrastructure

11

**Dell NativeEdge Platform**  
Edge Operations Simplified  
Infrastructure Management and Application Orchestration

12

**Comprehensive Edge and  
AI Services**  
Expertise in Design, Development, and Deployment  
Leveraging Dell Technologies' Services and Ecosystem

13

**About Dell Technologies  
OEM Solutions**  
A Trusted Partnership for Innovative Designs and  
Faster Time to Market  
Engineering Excellence, Program Management,  
Secure Supply Chain and OEM Services

# Design AI inferencing solutions for the edge: here & now.

Artificial intelligence (AI) has come a long way. For years, what we now call Traditional or Numerical AI development has primarily focused on core or cloud-based analytics that have helped organizations improve operations and better understand their customers. Through powerful data-driven insights, businesses have been able to predict trends, optimize processes, and deliver more personalized experiences.

But now, we've entered a new era in AI's evolution with Generative and Conversational AI.

Advances in computing technologies from companies like Dell Technologies and its partners enable solution builders to deploy powerful AI and GenAI models for inferencing at the edge,

enabling real-time data processing and decision making. This opens the doors to an entirely new set of possibilities and monetization opportunities for solution builders.

Whether in a factory, an intensive care unit, or a retail checkout, with Edge AI the future isn't just about insight, it's about action. Autonomous inferencing and AI-aided human decisions help people, devices, and machines act faster and smarter.

But what, specifically, can solution builders do with Edge AI inferencing?

And how can you make sure you deploy it into your solutions successfully?

**Let's explore.**



With Edge AI, the future isn't just about insight, it's about action.



# Inferencing at the Edge: The next phase of AI creates a trillion-dollar opportunity.

IDC predicts AI is poised to generate up to a trillion dollars in productivity gains.<sup>1</sup> And over 60% of data compute will reside in edge environments by 2027<sup>2</sup>. However, the future of AI isn't a binary choice between edge and cloud. The cloud or data center will serve as a training and customization environment for sophisticated and powerful AI models. Meanwhile, edge technologies will handle fast-paced inferencing.

With inferencing, organizations use a pre-trained AI model to make a prediction, an autonomous decision, or a real-time reply to achieve new levels of operational efficiency, responsiveness, control and customer experience on new sets of data.

1 IDC InfoBrief commissioned by Dell Technologies, Workforce Upskilling for the AI Era, January 2024  
2 451 Research Market Monitor analysis of low-latency workloads across eight key industries in the US; Edge Workload Total Addressable Market (TAM) Analysis, 2023.

## Cloud / Data Center AI

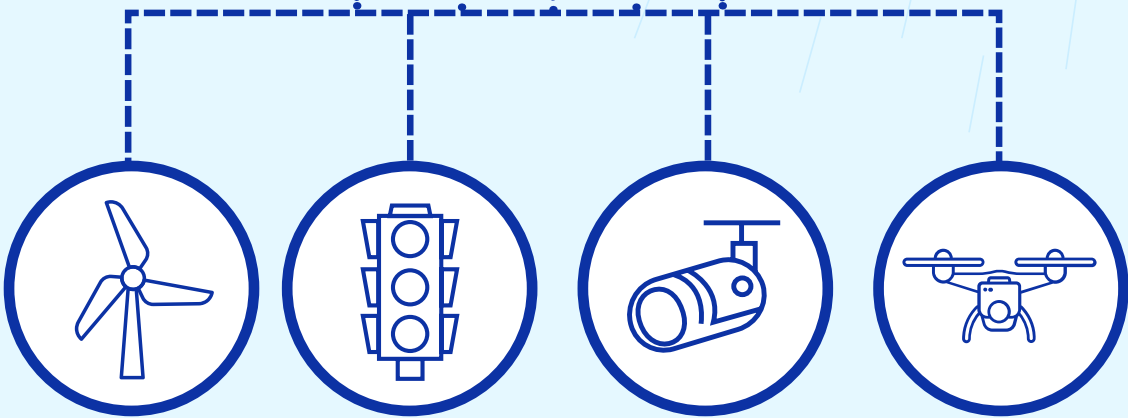


Compute | Training | Model  
Fine-tuning & Customization

## Edge AI



Edge Server



Compute | Inferencing | Decisions

Training Data

Distribute Trained Model

For solution builders, the ability to bring compute power right next to the data source, in almost any environmental location, while including zero touch and Zero Trust principles, creates compelling value. And Dell Technologies is here to help make it happen with AI infrastructure optimized for the edge, powered by NVIDIA, Intel, AMD and other partners.





# Why Edge AI? Why now?

Edge deployment models offer significant advantages:

## Reduce latency

Enable real-time decision making

- **Reduced Latency:** With AI at the edge, data is processed near its source, minimizing latency associated with sending data back and forth to the cloud or data center.
- **Immediate Impact:** This allows for real-time insights and immediate actions, crucial in industries such as manufacturing, healthcare, autonomous vehicles, and financial services, where split-second decisions matter.

## Enhance security

Minimize the risk of data breaches

- **Data Control:** With AI processing happening locally, sensitive data can be kept on the edge device, reducing the risk of exposure during transmission.
- **Stronger Security:** Keeping data on-site helps protect against potential breaches, while also ensuring that Zero Trust security principles and privacy regulations are more easily met.

## Control costs

Operate in environments with limited connectivity

- **Energy Efficiency:** By utilizing less power-intensive, and smaller-footprint devices at the edge, companies can significantly reduce energy usage, contributing to both operational savings and sustainability efforts.
- **Bandwidth and Networking:** Edge AI minimizes bandwidth usage by processing data locally on devices, saving costs and enabling operations in environments with limited connectivity.

By adopting AI at the edge, solution builders can deliver faster, more secure, and energy-efficient solutions that empower real-time insight and action. It's the next step in AI's evolution, and it's happening now.



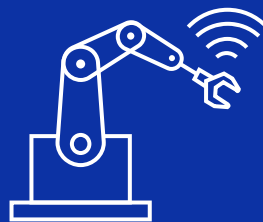
# Your industry. Your edge. Your AI.

Whether it's pinpointing a tumor on a CT scan and generating a report, detecting product defects on a factory assembly line, or recognizing the difference between parsley and cilantro at a grocery store checkout, Edge AI is poised to transform operations across industries.

For solution builders, concepts once considered science fiction are achievable now.

Layering AI and GenAI Edge use cases blend predictive tasks with creative solutions, fostering a holistic approach to addressing a wide range of scenarios. For instance, within the medical industry, Edge AI can help practitioners evaluate X-rays and MRI scans for anomalies, while GenAI writes the initial reports. Together, they streamline diagnostic processes and improve medical reporting efficiency.

## Manufacturing



- Product defect detection
- Smart inventory management
- Improved employee safety
- Robots and digital twins

## Computer Vision



- Crowd control and security
- Queue management
- AI-powered video analytics
- Facial recognition

## Distribution



- Drone fleet management
- Machine defect detection
- Fuel optimization
- Smart logistics

## Healthcare



- AI-aided CT-scan analysis
- Disease detection
- Detecting machine malfunctions
- Hospital resource optimization

## Retail



- Smart shelves (inventory)
- Shoplifting detection
- Interactive digital signage
- Dynamic pricing, real-time deals





# Putting AI into action at the edge.

Across industries, AI inferencing at the edge is already delivering tangible results. Together with Dell Technologies, OEM customers and innovative companies are utilizing AI to solve complex challenges at speed. From enhancing operational efficiency to improving patient care, these real-world customer success stories exemplify how AI inferencing at the edge delivers meaningful outcomes.



Dell Technologies is #1 for edge storage and edge servers.<sup>3</sup>



### AI fault detection at 125 mph.

Duos partnered with Dell Technologies and built the Railcar Inspection Portal (rip®) with Dell PowerEdge servers. The trackside data center runs AI algorithms at the edge that inspect 360-degree images of every car in a train as it travels through a rip®. PowerEdge servers also enable AI model training in a central data center.

**120X** performance boost using AI inspections

**8X** increase in inspection accuracy



### Formula 1 racing reimaged with AI.

McLaren Racing analyzes billions of data points on their cars daily utilizing a wide range of Dell high performance computing technology from PowerScale storage to PowerEdge servers and Precision mobile workstations. The racing team performs aerodynamic simulations and tests car enhancements via 3D digital twins that inform car design and part modifications.

**90%** reduction in time required to design and manufacturing of new car components

**300+** onboard sensors boost car performance and speed via real-time data analyses of 100,000+ parameters



### Disease diagnoses in under a minute.

Partnering with Dell Technologies OEM Solutions, Anke launched its AI-powered ANATOM S800 CT scanner months ahead of schedule. Powered by custom Dell workstations, the scanner creates ultra-high resolution and image processing technology based on AI to double diagnostic accuracy and enable quicker treatment pathways.

**100%** increase in diagnostic accuracy

**60** seconds or less to identification of conditions such as respiratory disease

<sup>3</sup>: IDC Worldwide Quarterly Enterprise Infrastructure Tracker: Buyer and Cloud Deployment, Q3 2024, January 2025. Based on revenue. Edge storage and edge servers refer to the heavy edge category, which IDC defines as the physical equipment that serves as the foundational infrastructure for edge.



# Simplify your journey to Edge AI inferencing with us.

79% of organizations believe AI has the potential to save time or money<sup>4</sup>. But AI is not without its challenges — especially at the edge. As solution builders innovate with AI, they need to address a range of considerations including:

- 

Talent and skill gaps
- 

Environmental factors
- 

Space constraints
- 

Industry certifications
- 

Security
- 

Scalability
- 

Infrastructure lifecycles

While these challenges are real, they're not insurmountable. With the Dell Technologies AI Factory methodology, you have access to the services, ecosystem partnerships, and infrastructure required to design and deploy your AI solutions faster, with higher quality, and with less risk.

<sup>4</sup> CIO Market Pulse White Paper: How edge can maximize resources and boost productivity.

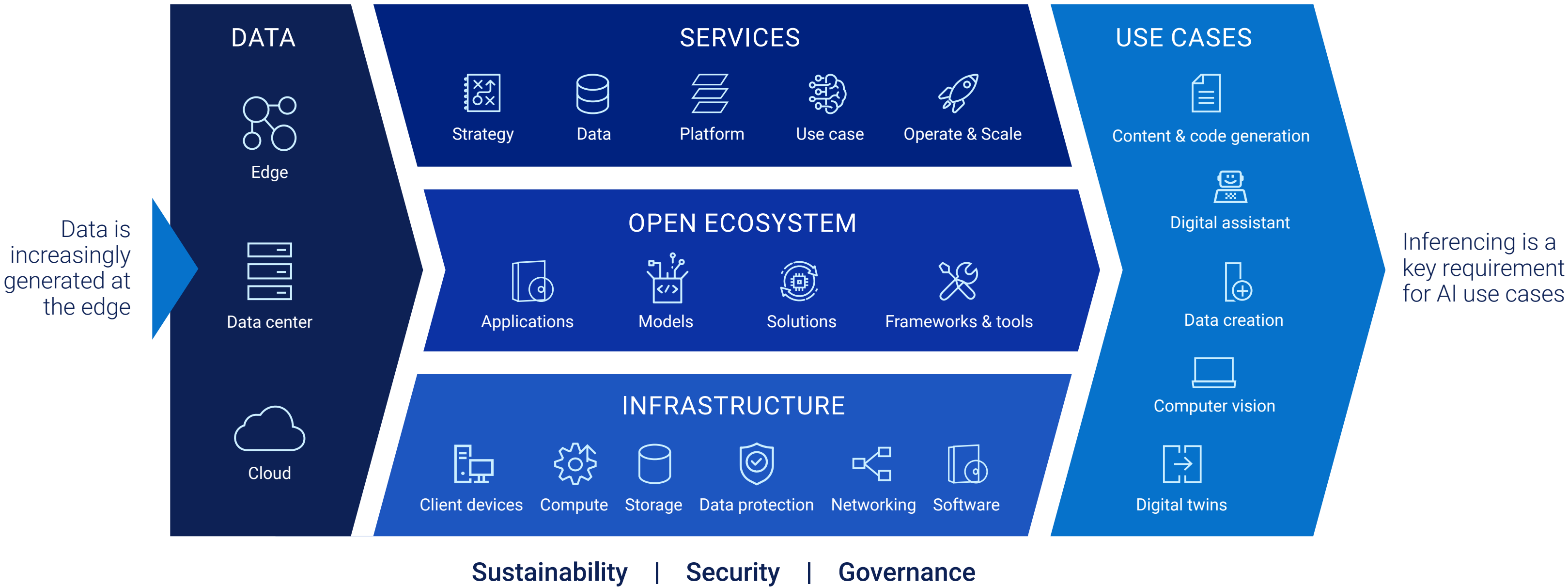




# Fast-track Edge AI innovation with the Dell AI Factory

The Dell AI Factory represents our strategy for embracing and implementing AI. AI factories power the AI revolution, generating actionable intelligence, fresh content, and new insights in place of physical goods.

Any AI journey should begin by understanding its destination: the use case. The Dell Technologies AI Factory is relevant to all AI types, use cases, and industries.



## Our strategy for making your AI use case real involves four elements:

### 1. Data

Data is the fuel for the AI Factory. We help you make the most of your and your customers' data, which is increasingly generated at the edge. Leveraging our expert AI services team, we can help you prepare data seamlessly to ingest to the AI platform, ensuring valid and impactful outputs.

### 2. Infrastructure

Leverage our platforms for AI including servers, storage, networking, and PCs that are optimized for edge conditions. Defeat temperature swings, dust, and humidity with infrastructure you can count on to deliver high availability and low latency in critical environments — from battlefields to farm fields.

### 3. Ecosystem

Harness the power of an open ecosystem, enabling diverse AI models, frameworks, and hardware for maximum flexibility. Dell and its technology partner network ensure you have the capabilities, solutions, and tools to deliver innovation faster, cut costs and increase value from data.

### 4. Services

We simplify and accelerate your Edge AI journey with our comprehensive suite of edge and AI services. By combining your unique expertise with ours, you'll give your customers the confidence that their Edge AI deployments deliver value more quickly and are optimized, secure, and resilient.



# Infrastructure is the foundation of the AI Factory.

Dell offers the broadest AI portfolio spanning from desktop to cloud<sup>5</sup>, allowing you to right-size your AI investment and giving you the flexibility to deploy AI anywhere. Edge AI infrastructure often requires specialized hardware that supports deployment in harsh edge environments. Our edge-optimized platforms for AI provide solution builders specific capabilities such as:

- Ruggedized platforms
- GPU-enabled systems
- Small form-factor products
- Long life systems
- Industry certifications
- OEM-Ready solutions (de-branded, re-brand ready)
- Customization capabilities

## Shorten time to value with Validated Designs for AI

Leverage our tested and proven configurations, designed from the start to fit customer needs based on specific AI use cases and help speed and simplify deployment of new solutions.

[Learn more about Dell Technologies AI Solutions](#)



### PowerEdge Servers

Fast-track your Edge AI goals using PowerEdge servers with superior acceleration, diverse GPU options and rugged, small footprint and industry certified platforms.



### PCs and Workstations

Allow AI processing locally on the device. Our broad portfolio of AI workstations and AI PCs provide the necessary hardware and software infrastructure to enable AI inferencing at the edge, empowering organizations to leverage the power of AI in real-time, even in resource-constrained environments.



### Edge Gateways

Connect OT/IT environments and extract value from edge-generated data with no interruptions to the infrastructure. Compact in design, they enable you to collect, consolidate, and perform lightweight analyses on vast Edge AI data.



### Storage and Data Protection

Help customers unlock the hidden value of their data and protect and secure their Edge AI workloads. Secure Edge AI workloads against data loss, corruption, and cyber threats with modern, simple, and resilient data protection solutions.



### Hyperconverged Infrastructure (HCI)

Benefit from the breadth of the Dell Technologies HCI portfolio that allows for choice based on your desired outcomes.



### Networking

Enable simplified design, management and monitoring of powerful Ethernet fabrics to handle modern workloads like Generative AI and Edge AI inferencing.



### Data Management

Empower your customers to access data across edge, core and multicloud to power analytics workloads like Edge AI and GenAI for faster model tuning and business insights.

<sup>5</sup> Based on Dell analysis, August 2023. Dell Technologies offers solutions to support AI workloads across 12 product and service categories



# Simplify and accelerate your Edge AI deployments with Dell NativeEdge

Dell NativeEdge is an edge operations software platform that securely scales infrastructure and orchestrates Edge AI applications across any location. This platform makes it easy to deploy and manage AI applications across many distributed endpoints and supports edge inferencing applications.

NativeEdge Blueprints further simplify development, deployment and testing for various GenAI and Edge AI use cases. These pre-tested and pre-validated designs represent full-stack, integrated hardware and software solutions — ready to be deployed at the edge on NativeEdge Endpoints, including servers, workstations, desktops, and gateways, powered by the latest Intel processors and NVIDIA GPUs.

[Learn more about Dell NativeEdge](#)



## Accelerate AI innovation at the Edge

**Infrastructure management**  
Centralize and streamline edge operations at scale using Dell's zero-touch, Zero Trust infrastructure management to manage hundreds or thousands of edge devices across multiple locations.

**Application orchestration**  
Effortlessly deploy, manage, and scale Edge AI applications across your customers' edge environments. Enable your customers to select vertical or use-case-specific ISV applications from a catalog of blueprints, and to choose cloud environments that best suit their edge workloads.

**Edge to Multicloud**  
Enable your customers from edge to core to cloud. With NativeEdge your customers can orchestrate their applications not only on Dell devices and infrastructure, but also on their existing infrastructure, as well as in multicloud environments with the help of virtualization and containerization.



# Provide a full range of Edge and AI services to your customers.

Navigating the complexities of AI solution design, development, and deployment at the edge requires not only cutting-edge technology, but also the right expertise. Many businesses lack the specialized knowledge needed to successfully bring AI solutions to production, creating a tremendous opportunity for organizations that can bring together a full complement of Edge AI services.

Empower your customers by leveraging Dell's broad and tailored suite of services for the edge and deliver your unique capabilities, such as integration, consulting, implementation, break/fix and managed services. Your services combined with Dell Technologies expertise will provide your customers assurance that their Edge AI solutions are optimally and securely deployed and supported.

With Dell Technologies and our ecosystem partners, including NVIDIA, Intel and AMD, you can leverage a world-class infrastructure, software and services and deep industry expertise to bring AI to the edge — quickly, efficiently, and with confidence. Together, we'll ensure your AI solutions are built for success, empowering you to stay ahead in this rapidly evolving landscape.





# AI Better Together – Your Opportunity

You're at the forefront of delivering AI-powered solutions that transform industries. Whether it's real-time diagnostics in healthcare, predictive maintenance in manufacturing, or smart retail experiences, your customers rely on you to bring cutting-edge AI to the edge.

Let us provide AI-optimized infrastructure, customized hardware, and seamless integration to help you deploy scalable, secure, and efficient Edge AI solutions—faster. With our proven technology, global supply chain, and deep expertise, you can accelerate time to market and focus on what you do best: building innovative solutions for your customers.

## OEM Customer Needs

- One-stop shop – from model building to edge deployment
- Fast access to the latest technologies and solutions
- Accelerate time to market
- Adapt to customers' evolving needs
- Create a robust AI and GenAI solution while offering your value-added services

## Dell Provides

- Portfolio for traditional and cloud/edge native deployments
- Relationships with NVIDIA, Intel, AMD among others in the partner ecosystem
- NativeEdge platform, industry Blueprints and AI Validated Designs to save you time and effort while reducing risk
- Right-sized deployments that better align to customers' strategies
- Augment your capabilities with our partner ecosystem & Dell Services





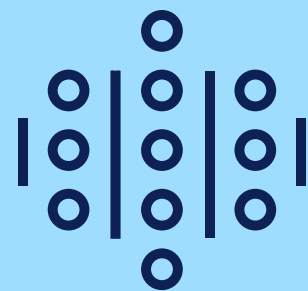
# Why Dell Technologies



The broadest AI portfolio at the edge



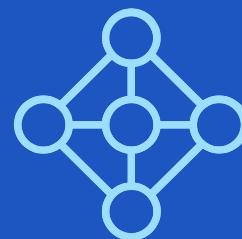
Dell NativeEdge  
Purpose-built Edge portfolio



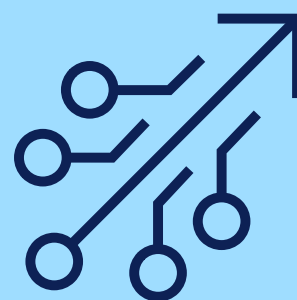
Accelerate AI innovation



#1 OEM solutions provider  
worldwide since 2016<sup>6</sup>  
Global ecosystem of  
technology partners



End-to-end AI and Edge  
solutions tailored to  
your customers' needs



Optimize for AI



Reliable global supply  
chain and support

We're always innovating so you can capitalize on AI insights at the edge.

6 Global Share, VDC Research, Worldwide OEM Solutions Provider.



# About Dell Technologies OEM Solutions

Dell Technologies OEM Solutions helps solution builders and product development organizations bring their products to market faster and with less risk by incorporating Dell products with their own Intellectual Property (IP).

OEM Solutions has a broad, innovative portfolio, unique capabilities such as design, engineering and program management, world-class services and a reliable, resilient and secure supply chain. We have specialists dedicated to OEM business comprising Product Group, Engineering, Program Management, Procurement, Sales, Marketing and more, as well as a broad partner ecosystem — and we’ve been doing it for 25 years.

[Learn more at Dell.com/OEM/Edge](https://www.dell.com/oem/edge)

Copyright ©2025 Dell Inc.. All Rights Reserved. Dell Technologies, Dell and other trademarks are trademarks of Dell Inc. or its subsidiaries. Other trademarks may be trademarks of their respective owners.

Our OEM Solutions team simplifies every stage of your product lifecycle.

