



Smarter AI starts with better data

Unlock the potential of AI with a robust data pipeline.



The AI revolution will touch everything

The evolution of civilization is marked by transformative milestones, each reshaping how humanity lives, works, and connects. The transition from nomadic lifestyles to organized settlements allowed for the development of cities, fundamentally altering human existence.

These early cities became hubs of innovation, trade, and cultural exchange—driving advancements in transportation, commerce, governance, and social structures.

Today, it's AI that will become the next fundamental turning point in our journey.



The AI Data Pipeline

Data needs unified search, meta-tagging, cataloging, resilient storage, and compute horsepower.

AI learns from data and prepares it for models capable of making decisions, predicting trends, and generating content. The more data you feed AI, the smarter it becomes.

While it can be complex to make sense of structured and unstructured data, companies like Dell and NVIDIA are making it easy to harness the power of AI using a robust pipeline that handles everything from data collection to model deployment.

Data Management

AI/Analytics

Security and Governance

NVIDIA Accelerated Computing Platform

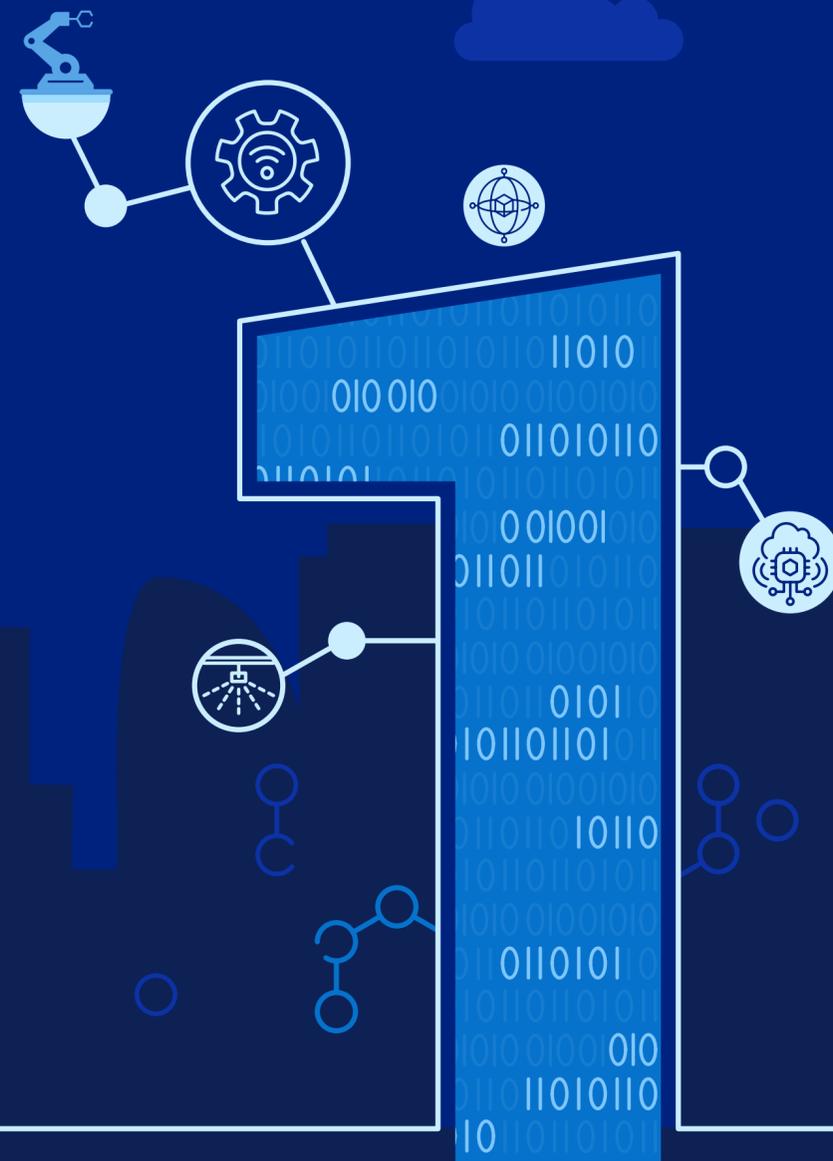
, which includes AI infrastructure and AI software and models, underpins the computational power and software tools required for this work. At the same time, Dell PowerScale provides the secure infrastructure needed to store and process massive datasets.

Together, they have a successful AI Data Pipeline that watches the data all the way through the process to achieve effective outcomes—enabling businesses to unlock the potential of AI.

Data Creation

Applications | Human Machine Tools | Sensors

Much like humans have sensory abilities, our environment has sensors at the edge that collect data requiring measurement, calibration, and processing. This data creation is the lifeblood of the AI pipeline. All data types, from health to proximity, motion and vision, have their place in driving AI outcomes.



Watch a video about data creation methods using NVIDIA edge solutions and Dell infrastructure >

Start a foundation for accurate reliable outcomes.

The way you gather, process, store, and manage raw data leads to valuable insights for refining models, identifying biases, and ensuring data quality.



Watch a video about data ingestion methods using NVIDIA edge solutions and Dell infrastructure >

Power AI with the right data in the right format.

Effective data ingestion requires robust mechanisms to handle diverse data types, volumes, and formats while ensuring data quality, consistency, and compliance with regulatory standards.



Data Ingestion

Replication | Workflows | Streaming

Data ingestion requires modern technology to ensure raw data, whether structured or unstructured, is seamlessly integrated into the pipeline in a format suitable for downstream tasks such as transformation, training, and model evaluation, fine-tuning and inference. This stage is critical in building trustworthy AI systems that perform reliably and ethically. Data security and governance prevent privacy breaches, ensure compliance, maintain data quality and uphold fairness.

Data Aggregation

Storage | Formatting | Cleansing

Raw data is messy.

It may contain errors, duplicates, or irrelevant information. For example, if you're processing customer reviews, duplicate entries must be removed, typos corrected, and data categorized. Anyone who collects data becomes a data product owner, and anyone who produces the data is packaging and storing it safely, securely, and reliably.

Data aggregation involves cleaning, organizing, and preparing the data into structured formats suitable for training, analysis, or deploying AI models. Data owners and producers must trust that the data has integrity and accuracy. And that it is compliant and secure to mitigate privacy and cyber security concerns.



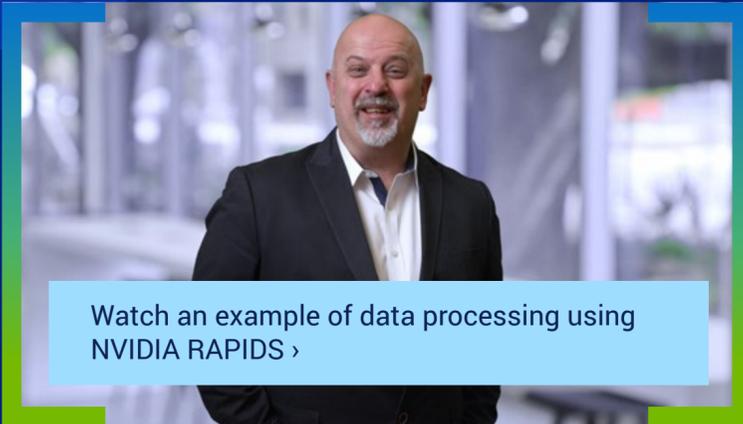
Build a scalable, reliable, secure home for data.

Dell PowerScale ensures that massive amounts of data can be efficiently housed while adhering to stringent privacy and security regulations.

Data Processing

SQL | Spark | ML/DL | Indexing

When data is pulled together during the **aggregation stage**, it's not ready for consumption; it's traditionally analyzed and propagated across environments, such as AWS, which can take months. But there's an effective way to turn months into moments: utilizing a data lakehouse. This pulls the data together in real time and uses memory to persist that data within a single environment so it can be fed to dashboards and applications.



Turn months into moments.

Did you know data hosted on an existing PowerScale ecosystem can use native replication and protection features extended in Microsoft Azure? The platform supports GPU-accelerated Apache Spark engine using NVIDIA RAPIDS, included with NVIDIA AI Enterprise, enabling efficient real time insights without any code changes.

Model Training and Refinement

Exploration | Metrics | Inference | Tuning

By feeding processed data into models, AI begins to recognize patterns and make predictions, using that data to train, validate and refine models, and to augment inference. The errors between predictions and actual results are measured and optimized to adjust the model's parameters, leading to an assessment of the final model's performance. It's an iterative approach that can use a tremendous amount of processing power.



Get on the fast track for successful data outcomes.

NVIDIA's accelerated computing platform and Dell PowerEdge drastically reduce the time needed for training, allowing businesses to iterate faster.



Watch a video about real-time performance tracking >

Smoothly integrate into the flow of business.

Dell ensures seamless deployment with infrastructure that supports scalability and reliability, even with intensive data workloads.



Data Outcomes

Content & Code Generation | Digital Assistants | Data Creation | Computer Vision | Digital Twins

Once the model is trained it must be integrated into real-world applications. This could mean using AI inference to deploy edge applications, websites, digital assistants, or operational systems. To achieve the desired outcome, the data flow must be correct, and the result sets must be accurate. The outcomes are stored, processed, and preserved—providing more data to ensure that the entire system functions correctly, that data is processed effectively, and that the AI models perform well over time.

Data Security and Governance

Data controls | Model controls

Security is non-negotiable.

As data flows through various stages of the pipeline, it's essential to protect information from unauthorized access, breaches, and misuse. Implementing enterprise-grade encryption, consistently monitoring, controlling access, and securely authenticating helps safeguard data integrity and confidentiality.

Furthermore, compliance with relevant regulations, such as GDPR, HIPAA, or CCPA, ensures that the pipeline adheres to legal and ethical standards for data privacy. By integrating data security and compliance measures into the pipeline, organizations can build trust among stakeholders, creating the basis for scalable and responsible AI solutions that align with organizational goals and the expectations of society.



Protect data from start to end.

Dell PowerScale is the world's most secure storage, ensuring that data is always protected and safe from breaches or cyber attacks.

Dell AI Data Platform

Open | Flexible | Secure

Designed for the new AI era to place, process and protect data. It is built on an open, flexible and secure architecture that caters to the dynamic demands of Data and AI teams.

PLACE



Enable effective data placement to ingest large volumes of data from various edge devices in real-time, aggregating data from multiple sources on petabyte scale persistent storage and performance for GPUs at scale.

PROCESS



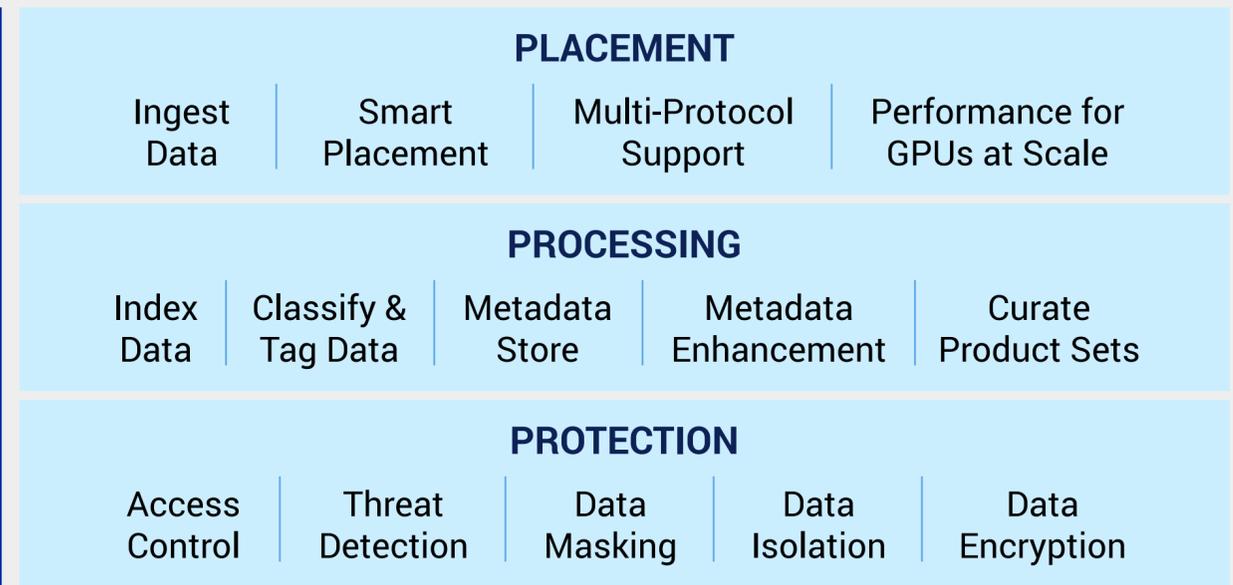
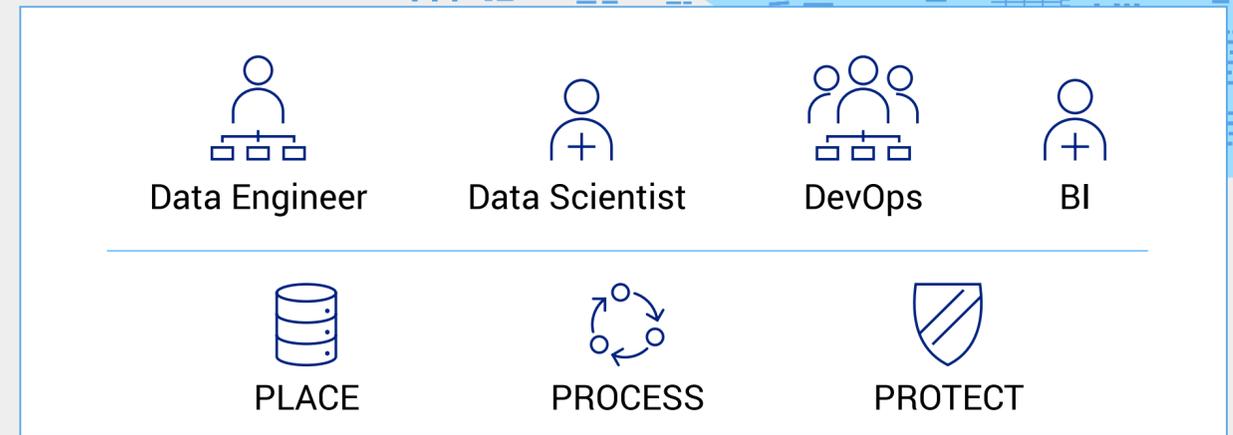
Unlock valuable insights with quality, enriched unstructured data combined with structured data enabling the curation of comprehensive product sets to drive advanced analytics and AI models.

PROTECT



Operate confidently and safely using a comprehensive, proactive data protection approach that ensures your data remains protected, resilient and compliant.

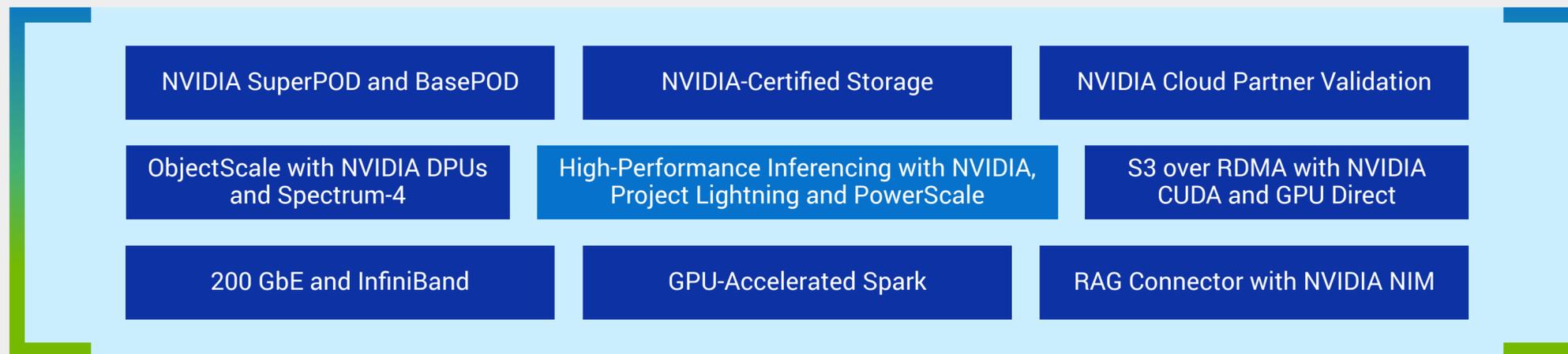
DELL AI DATA PLATFORM



Dell AI Data Platform

In partnership with NVIDIA

Together, Dell Technologies and NVIDIA are redefining what's possible for AI-driven enterprises with the Dell AI Data Platform. This end-to-end solution integrates NVIDIA's revolutionary AI hardware and software with Dell's cutting-edge infrastructure, delivering unmatched scalability, performance and reliability. By combining validated and tested architectures, this collaboration removes the complexity of full-stack AI data center deployments, empowering businesses to adopt AI seamlessly, optimize their operations and accelerate innovation. From enabling real-time insights to powering advanced AI models, this partnership unlocks tangible value and strengthens organizations' abilities to lead with confidence in an AI-first world.



Explore how Dell AI Factory [can help you drive your AI strategy forward](#)

Read the 7 Steps to Unlocking [the Value of Data for AI](#)

Learn more about the [Dell AI Data Platform](#)

Copyright © 2025 Dell Inc. or its subsidiaries. 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. Intel® Xeon® is a trademark of Intel® Corporation or its subsidiaries.