Data is the DNA of AI

How data fuels competitive advantage and better AI business outcomes

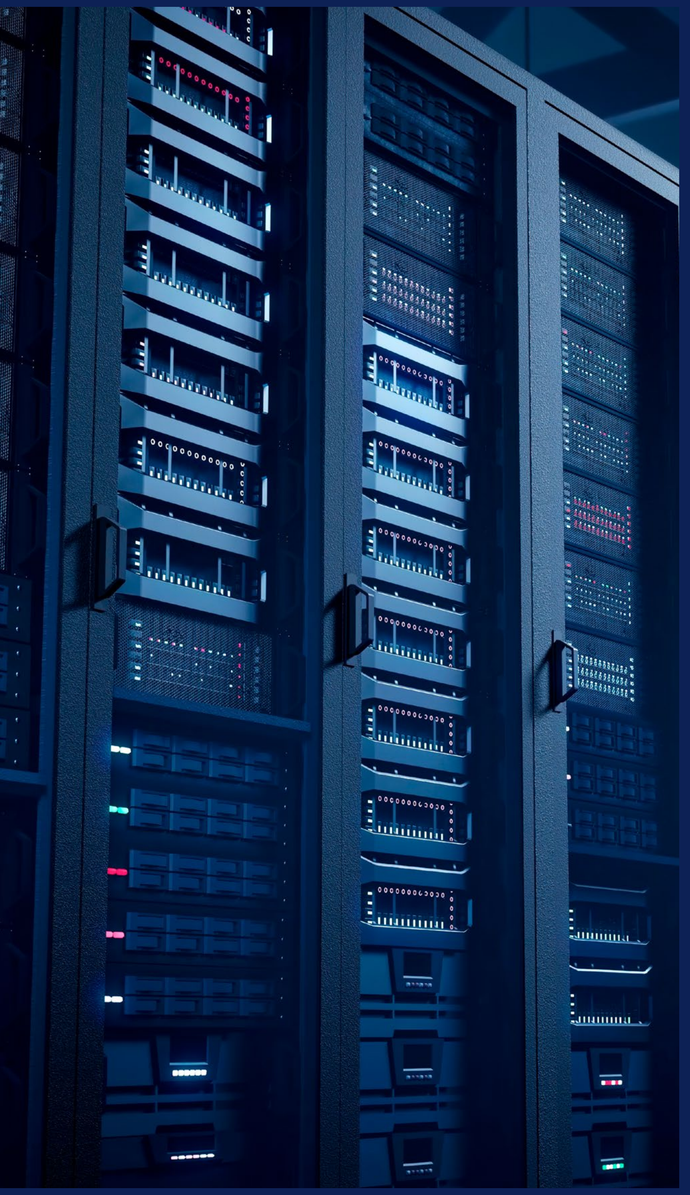
Data is the DNA of AI

For years, the familiar metaphor “data is the new oil” has shaped how organizations approached digital strategy, highlighting the value of data and the need to process and refine it. But in the age of artificial intelligence, that analogy has outlived its relevance, and doesn’t accurately describe the immense value data provides. AI doesn’t just run on data like an engine runs on fuel. It learns from data. It reasons through data. Data drives the context, understanding, and outcomes of what’s possible with AI.

Data being the DNA of AI is a more appropriate analogy. DNA defines the traits, structure, and behavior of living things. For AI, data essentially fulfills the same role:

- Data determines how AI perceives, understands, predicts, and acts.
- Clean, connected, and well-governed data enables AI to perform with precision and flexibility.
- Inaccurate or fragmented data leads to inferior capabilities or unreliable outcomes.
- Data is not an output of operations. It’s an input to intelligence.

Understanding this shift is essential for building successful AI initiatives and driving innovation. Organizations should consider their current data foundation, start planning and preparing their data for the new world of evolving AI, and begin reimagining how data can fundamentally change how AI drives competitive advantage.



The best data for AI, and why it matters

While AI models often get the spotlight, it's the underlying data that determines how AI understands context, objectives, the role it plays, and how it adapts to new scenarios. Flawed data leads to flawed outputs, limited reasoning, and fragmented performance. Structured, contextual, and dependable data, by contrast, enables AI to deliver results and insights that are aligned with business goals.

Even the most advanced AI systems cannot deliver value if the data behind them is flawed. Consider the following data-related red flags to identify and address key data concerns that could undermine accuracy, performance, or trust.

Inconsistent, mismanaged, or data lacking context

AI systems require data that is clean, organized, and aligned across sources. When foundational elements are missing or inconsistent, AI performance suffers.

- Poorly handled data slows model training and increases reliance on manual intervention.
- Misaligned formats across systems make it harder for AI to detect patterns or produce accurate outputs.
- Data without context leads to surface-level insights, limiting personalization and relevance.

Unverified or untrusted data sources

Data quality and governance are essential for building AI that's accurate, compliant, and trustworthy. When data cannot be verified or traced, it introduces risk.

- Unattributed data lacks visibility into its origin, making it difficult to assess credibility.
- Biased or incomplete data can produce misleading results and reinforce flawed assumptions.
- Lack of governance creates uncertainty around data safety, privacy, and regulatory compliance.

Harness your data to drive competitive advantage

AI has the potential to reshape how decisions are made, how customers are served, and how value is delivered. For many organizations, the current state of their data is not a barrier, but an opportunity to build a stronger foundation for competitive advantage. By focusing on the following areas, it's possible to turn complexity into capabilities.

Fragmentation across systems and teams
Scattered data across departments, applications, and platforms presents an opportunity to unify and align. Consolidating data sources enables a more complete view of the business, reduces duplication, and gives AI access to deeper, more actionable insight.

Inconsistent quality and structure
Unstructured or outdated data can be transformed into a strategic asset. Investing in data cleaning, labeling, and standardization improves model performance and reduces the time and cost of AI deployment-unlocking faster, more accurate outcomes.

Limited trust, governance, and compliance
Modern businesses must remain in complete control of their data. Establishing clear data lineage, security controls, and compliance frameworks builds trust in AI systems. Maintaining a strong data governance structure and data audit trail turns risk into reliability and ensures sensitive data can be used confidently, responsibly, and with attribution for compliance.

Infrastructure that doesn't scale
Legacy systems highlight the need for modern, flexible infrastructure. Upgrading to platforms that support real-time access, scalability, and integration prepares the organization to meet increasing AI demands with speed and efficiency.

Proactively addressing these data considerations enables organizations to build a lasting foundation for AI innovation, agility, and growth, rather than simply accelerating AI adoption.

How to prepare your data for AI

Preparing data for AI is an essential step in realizing value from intelligent systems. It requires more than just storage or access. It calls for a purposeful data quality, governance, and infrastructure approach. You need proven frameworks to simplify complexity, accelerate outcomes, and support AI at scale.

Data quality

Accelerating AI outcomes starts with clean, well-labeled, and accessible data. High-quality input reduces the need for manual preparation and improves confidence in outputs. Establishing organizational standards for data structure, consistency, and completeness helps ensure the models your business relies on are learning from the right signals.

Protect and secure your data

AI should operate where the data resides, whether in the data center, at the edge, or in the cloud, while also fortifying cyber resilience through strict security, privacy, and compliance enforcement. Your organization also needs to implement a business continuity and data resilience plan.

Seamless and simplified data for AI

Managing AI workloads efficiently requires seamless data access and orchestration. Simplifying how data moves, how it's processed, and how it's governed ensures systems can scale without adding complexity. A unified, flexible approach enables faster experimentation, better performance, and long-term adaptability.

Bring your AI vision to reality

Achieving real outcomes with AI isn't simply about ambition, technology, or tools – it's about execution. The difference lies in how well your data can support the shift from isolated experiments to scalable, business-wide impact.

Dell Technologies and NVIDIA can help you leverage AI to drive innovation and achieve your business goals. The Dell AI Factory with NVIDIA is the industry's first and only end-to-end enterprise AI solution,[1] designed to speed 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.

- **Implement scalable AI infrastructure**
  The Dell AI Factory with NVIDIA combines Dell's enterprise-grade infrastructure with NVIDIA's advanced AI software and hardware to support training, fine-tuning, and model deployment.
- **Build a strong data foundation**
  Prioritize optimally located, high-quality, aggregated, and secure data to serve as a foundation and core driver of AI effectiveness.
- **Prioritize the right use cases**
  Focus on AI use cases that align with business goals and are supported by available, trusted data to drive real outcomes.
- **Modernize your infrastructure for AI**
  Evaluate whether your current infrastructure can support AI at scale, from training to deployment, across cloud, core, and edge.
- **Embrace an open, secure, and flexible ecosystem**
  Adopt interoperable architectures, standards, and file formats that integrate with your environment and support evolving AI needs.
- **Leverage expert guidance**
  Accelerate success by partnering with experts who can guide strategy, streamline deployment, and optimize performance at every stage.

1. Based on Dell analysis, July 2024. Dell offers solutions with NVIDIA hardware and software engineered to support AI workloads from PCs with AI-powered features and workstations to servers for high-performance computing, data storage, cloud native software-defined infrastructure, networking switches, data protection, HCI and Services.

Preparing your data for what’s to come

AI is advancing quickly, but the biggest differentiator will not be the speed of innovation but rather the speed of readiness. New AI capabilities are already part of enterprise environments, and organizations that focus on data readiness will be better positioned to adopt, adapt, and scale AI innovations.

Generative AI integration
The quality, context, and freshness of data directly influence how generative AI models produce responses. Inferior data leads to flawed outputs.

Agentic AI
Such systems rely on continuous feedback and structured data to make decisions, adapt over time, and operate with a greater degree of autonomy.

Multimodal AI
Works with labeled datasets across formats like text, audio, and images to interpret and respond in more human-like, context-aware ways.

Federated data models
These models train on decentralized, local data that must be consistently formatted, privacy-compliant, and securely governed across locations.

Democratize data access and discovery
Enable teams with curated structured, semi-structured, and unstructured data for improved discoverability and self-service to create new data sets and products.

Build to be future-ready
Organizations with a unified and well-managed data foundation are better equipped to adopt emerging AI tools and support open formats like Iceberg and Delta Lake.

Make data an integral part of your competitive DNA

As AI moves from experimentation and proofs of concept to implementation, the difference between hype and impact will come down to one thing: the quality of the data behind it. Treating data as the DNA of AI means thinking beyond infrastructure and focusing on intent. What business outcomes do you want AI to drive? What data types and sources will best inform your AI to drive meaningful and trusted results? You can prepare your organization by fully investing in data discoverability, enrichment, and insights. This will allow you to turn your data into a competitive advantage as you embark on your AI initiatives.

As you consider these concepts, keep in mind that it's not just about keeping up. It's about building a foundation that AI can learn from, grow with, and continuously deliver value through.

Data that's unique to your organization

Customer interaction data
Chat logs, emails, and clickstreams used for personalization and predictive engagement.

Transactional data
Purchasing trends and seasonal data used to drive forecasting and revenue insights.

Operational data
Logs and sensor data used to streamline processes and detect issues.

Unstructured data
Documents, recordings, and presentations turned into searchable intelligence.

Real-time data streams
Live inputs from apps and devices enabling fast, adaptive decision-making.

84%

of organizations agree it is important to incorporate their own enterprise data to support generative AI.[2]

2. Enterprise Strategy Group Research Report "The State of the Generative AI Market: Widespread Transformation Continues." September 10, 2024.

Getting started

Ensure you're on the right path and work with Dell and NVIDIA to get the right tools, infrastructure, and expertise to prepare your organization for every phase of your AI journey.

Preparing for AI starts with data. But staying competitive requires action. Learn more about how you can get started with AI and contact your Dell representative to set up a fee-waived Dell Accelerator Workshop for GenAI where you can get expert guidance on making your data AI-ready and building your AI strategy.

Get started with AI

Explore the Accelerator Workshop

Dell **AI Factory** WITH NVIDIA

DELL Technologies | NVIDIA