DELLTechnologies

Accelerate Your Business With Artificial Intelligence
Modern Infrastructure Is Key

Best Practices
for Deploying AI Workloads

DELL
Technologies

Accelerate the Business Value of AI

AS ARTIFICIAL INTELLIGENCE (AI) EMERGES from research projects to full production, it is fundamentally changing the way companies innovate and deliver services to their customers. Machine learning (ML)—a subset of AI—is particularly important for turning massive amounts of data into useful insights and making organizations more efficient, innovative, and competitive. Organizations that are adopting AI are finding that they can enrich customer experiences, increase revenues, and optimize operational efficiencies.

AI leaders drive more than 10% of revenue from AI.[1]

The path to successful AI implementations varies, but the most mature AI organizations share some common characteristics. **ESG surveyed** 750 strategic IT decision-makers to measure the relationship between AI success and modern infrastructure.

ESG research discovered that organizations farther along in their AI journey are succeeding with the use of modern servers with extensive automation capabilities, accelerator technology, and converged/ hyperconverged infrastructure. ESG credits the use of modern technologies for the reason AI leaders drive more than 10% of revenue from AI.[1]

These findings lead to a common question: How does an organization get to this level of AI maturity?

Today, the sound approach is to start with flexible, AI-optimized compute and storage plus data protection, expanding into accelerators and converged/hyperconverged infrastructure as you mature—in other words, leverage a modern data analytics platform. Dell Technologies' solutions and Consulting Services help organizations develop their modern data analytics platforms, facilitate streamlined workflows between lines of business, and unlock their data's value–all while solving complex business problems and driving outcomes.

Best Practices for Achieving AI Maturity

AI-optimized Compute

Scale easily to accommodate larger and more diverse data sets over time

Scalable Storage

Design and deploy an optimum storage foundation for all of your AI data

Comprehensive Data Protection

Protect data through the entire pipeline, across private, public, and edge clouds
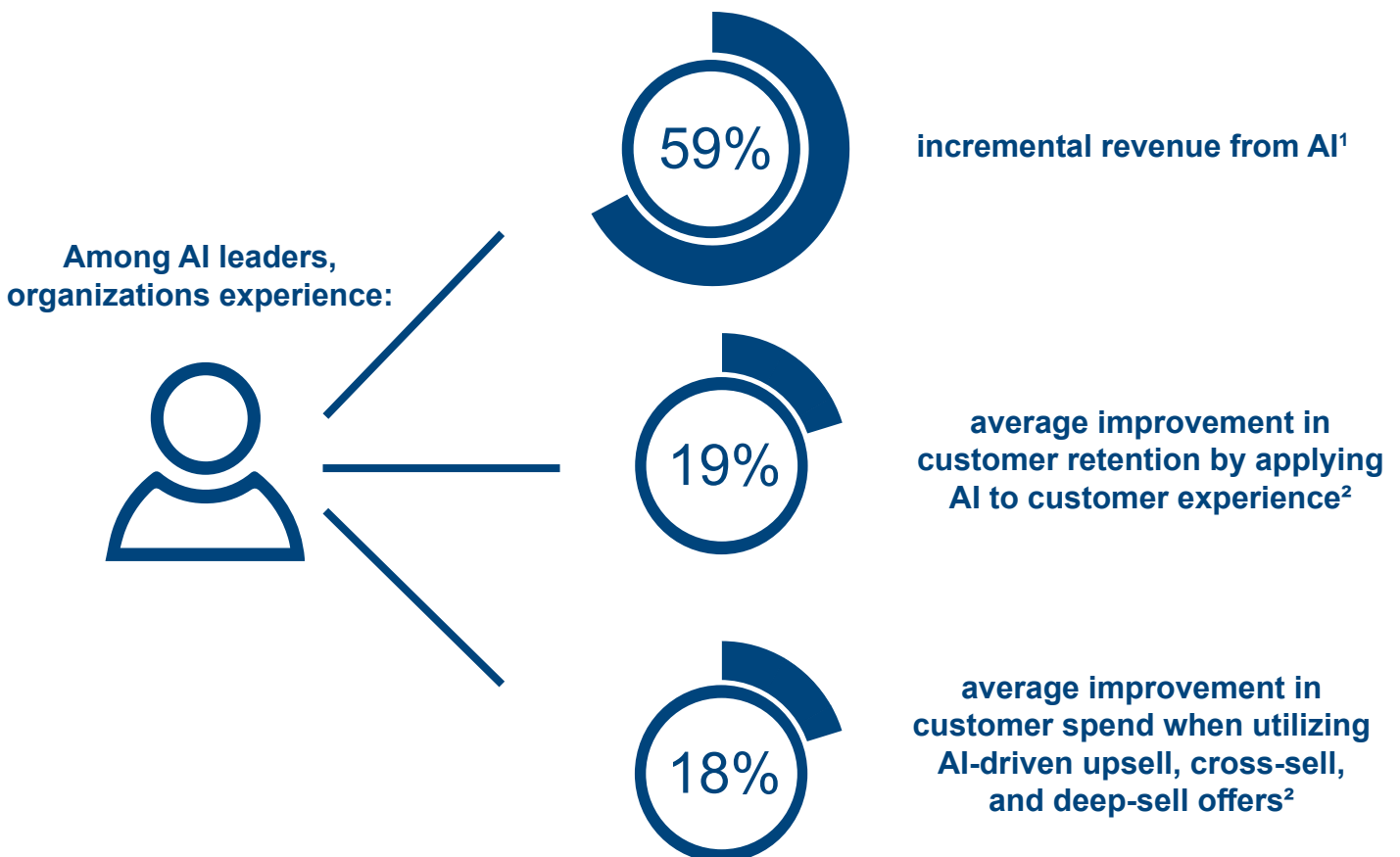
Consulting Services

Define, develop, adopt, and accelerate analytics and AI capabilities with the help of seasoned consultants

Accelerate the Business Value of AI

Identifying the Source of Value in AI Initiatives

Some organizations are having difficulty identifying the value from AI initiatives and the data analytics cycle. Collecting, integrating, and analyzing large volumes of diverse data is what ultimately leads to value through the insights unlocked from data. AI leaders are 7.8X more likely to drive significant business value with AI and 2X as likely to achieve faster than expected time-to-value with AI.[2] In fact, AI leaders are finding value across a range of metrics, including customer spend, customer retention, faster decision-making, improved quality, reduced IT downtime, and cost reductions due to automation.

ESG found there are multiple ways that AI leads to these improvements, such as freeing people from mundane tasks so that they can focus on higher value initiatives, identifying up-selling and cross-selling opportunities, and using sensor data to monitor, predict, and prevent production downtime. AI leaders are unlocking previously unrecognized value in data—59% of organizations see incremental revenue from AI.[1]
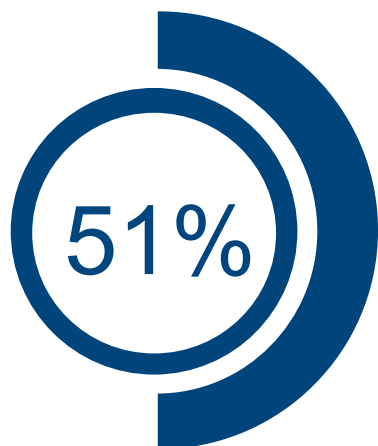
Among AI leaders, organizations experience:

59% incremental revenue from AI[1]

19% average improvement in customer retention by applying AI to customer experience[2]

18% average improvement in customer spend when utilizing AI-driven upsell, cross-sell, and deep-sell offers[2]

Grow Capabilities With AI-optimized Compute

AI WORKLOADS CAN BE COMPUTE-INTENSIVE. PowerEdge servers deliver the capabilities needed and can scale easily to accommodate larger and more diverse data sets over time.

For those just starting to deploy AI workloads, often the most practical and efficient way to begin is to leverage modern server resources. The 2nd Generation Intel® Xeon® Scalable processors with Intel Deep Learning Boost enable nearly 3X faster inferencing for deep neural networks.[3] Access to these AI-optimized servers is key to realizing the benefits of AI initiatives.

Leverage Modern Server Resources That Grow With You

51%

51% of AI leaders extensively use hardware accelerators for AI[1]

ESG research found a strong correlation between leveraging AI-optimized compute resources and a significant improvement in model development, as well as model use in production.

While any organization can begin its AI projects with existing modern servers, some AI use cases can benefit from accelerators, such as graphics processing units (GPUs) and field programmable gate arrays (FPGAs). While an organization may not need accelerators initially, 51% of AI leaders leverage hardware accelerators extensively for their AI workloads.[1] The need for acceleration depends on the use case or workload, the volume of data that needs to be processed, and the required time-to-result. PowerEdge servers come with built-in AI acceleration with 2nd generation Intel® Xeon® Scalable processors, as well as a variety of other accelerator options that can speed your AI workloads from analytics to training and inferencing.

Meanwhile, Dell EMC OpenManage automates the deployment, maintenance, and management of servers. This frees IT staff from routine management tasks and helps reduce time-to-result during incidents. With this kind of intelligent automation, servers can update themselves and maximize uptime.

PowerEdge servers deliver the high-speed processing performance needed for analytics and AI, as well as fast direct-attached storage and intelligent automation. In order to protect your data, PowerEdge has a cyber-resilient architecture, with hardware-enabled security capabilities directly on Intel® Xeon® processors and built into the firmware and hardware. PowerEdge servers also have modular scalability, which means they are designed to grow with your data and workloads over time.

Scale With Flexible Storage

THE DELL TECHNOLOGIES STORAGE PORTFOLIO is uniquely suited to the successful deployment of an IT infrastructure for AI. Whether you require the data lake consolidation of your unstructured data with Isilon, or the geo-distributed cloud object storage of ECS, the broad set of offerings will allow you to design and deploy an optimum storage foundation for all of your AI data.

For example, the Isilon F800 All-Flash Scale-out NAS® delivers the analytics performance and extreme concurrency at petabyte scale for AI. Isilon storage systems are designed to support three essential requirements of AI workloads: high performance, extreme scale, and data governance. A complete, robust AI infrastructure built on Isilon is an ideal solution to power and feed your organization's most demanding data-hungry algorithms.

High-performance storage is essential to continually deliver data fast enough to keep compute resources supplied with data. High latency for most use cases wastes compute capacity. The Isilon All-Flash storage systems deliver the throughput, IOPS, and capacity required to avoid I/O bottlenecks that slow model building and inferencing. Also, by using Intel® Optane™ DC memory and solid-state drives (SSDs) as temporary storage, organizations can accelerate real-time inferencing. Additionally, deploying Intel® 3D NAND SSDs for consolidated storage can reduce operating costs when compared to legacy storage systems.

The ability to scale storage is another key factor to meeting the needs of AI workloads. Storage needs to scale independent of compute resources. Organizations are amassing large volumes of historical data in data lakes and this is an ideal use case for Isilon storage systems that can scale up to 68 PB within a single cluster.

Isilon Storage System Designed to Support 3 Essentials

High Performance

The Isilon All-Flash storage systems deliver the throughput, IOPS, and capacity required to avoid I/O botlenecks

Extreme Scale

Isilon's automatic tiering capabilities enable organizations to deploy a single storage cluster for multiple storage use cases

Data Governance

Isilon's enterprise features include data management, data security, and data compliance

Scalable storage is a vital element to improving the quality of AI models. The more data is used to train an AI model, the more effective that model will be. In fact, ESG found that "all else being equal, an AI model trained with 1 petabyte (PB) of data will be more effective than a model trained by 100 terabytes (TB) of data, which will be more effective than a model trained by 1 TB of data."[1] However, as your organization's data ages and increasing amounts of active data enter your AI workflows, you'll need a solution to tier that data for backup and archival. Dell EMC ECS meets this need by providing a platform to consolidate archives and reduce demand on your primary storage. With Dell EMC ECS, you can easily move inactive older data to a low-cost storage tier and significantly increase your primary storage capacity. By harnessing the power of enterprise-grade object storage, you can streamline your storage infrastructure and power next-generation solutions.

Organizations should be able to choose price, capacity, and performance characteristics for different kinds of storage. ML and data analytics workloads need high I/O, while archives should utilize high-density storage. Isilon's automatic tiering capabilities enable organizations to deploy a single storage cluster that meets the needs of multiple storage use cases. Isilon's enterprise features for data management, data security, and data compliance, coupled with Dell Technologies' data protection, help your AI solutions conform to regulatory and enterprise security policy requirements.

Ensure Integrity With Data Protection

GIVEN THE MASSIVE AMOUNT OF DATA being used to train and feed AI models, data has become many organizations' most important asset. This data must be protected throughout the entire lifecycle, across private, public, and edge clouds.
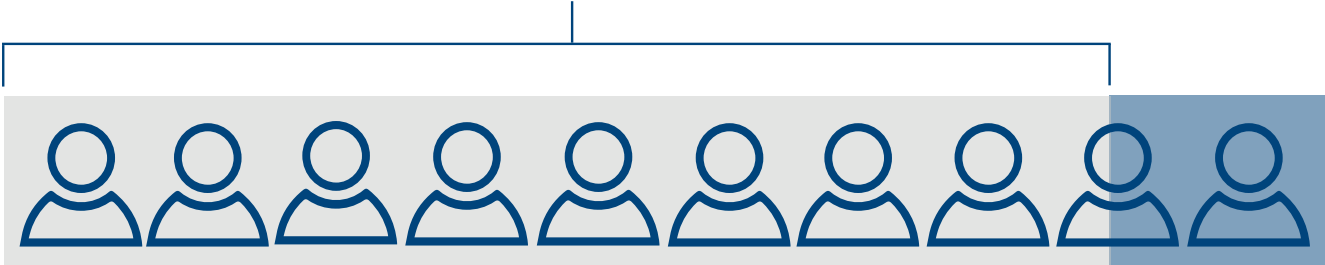
Dell Technologies data protection has proven itself in the industry with its high-performing, efficient data transport and commitment to data integrity—two key elements to meeting today's data protection demands. One critical component of the Dell Technologies data protection portfolio is the PowerProtect offering, which includes PowerProtect DD along with PowerProtect Data Manager. Leveraging both highly efficient data protection storage along with intelligent policy management and orchestration, PowerProtect allows organizations to keep AI workflows running and adds consistency to AI workflow processing.

Business systems—including AI systems—are becoming more and more vulnerable to cyber incidents and destruction by bad actors through the use of malicious tools, such as malware or ransomware. Building up the resiliency that is required to face these threats—along with the protection from operational and technology disasters—is paramount for today's successful organizations. Therefore, data isolation polices and automated recovery processes should be included in any data protection plan. Dell EMC's Cyber Recovery complements existing data protection paradigms by providing additional layers of protection against cyber incidents.

Data integrity ensures availability.

AI leaders are distinguished by their understanding of the need for data protection—they're 3X more likely to employ comprehensive AI data pipeline protection, with 84% of AI leaders making data protection a priority.[1] Dell Technologies' data protection solutions, made up of highly efficient data protection, storage, and intelligent orchestration software, assure broad support and compatibility to integrate with leading AI applications, along with the ability to bring data protection efficiency to customer applications.

84% of AI leaders make data protection a priority[1]

Unleash Innovation With Consulting Services

RAPIDLY SCALING AN ORGANIZATIONS' AI CAPABILITIES is challenging—and there are many opportunities for blind spots and costly mistakes. The **Dell Technologies Consulting Services** team of seasoned consultants helps organizations of all sizes, industries, and maturity levels define, develop, adopt, and accelerate their analytics and AI capabilities—from strategy through execution and data modeling, to enabling and optimization.

ESG found that AI leaders are 2X as likely to deeply leverage third-party services for their AI needs.[1] The Dell Technologies Consulting Services team engages with customers who would like guidance or partnership in determining how to get started, where to go next, how to develop alignment, or how to achieve optimal value from analytics investments. Dell Technologies Deployment Services offerings bring your new IT investments online as quickly as possible, accelerating adoption while freeing your IT staff for more strategic work.

Dell Technologies offers a full suite of services across all the stages of AI and advanced analytics needs:

- **ProConsult Advisory services** facilitate beneficial and lasting change, and focus on modernizing analytical platforms and architectures for scalability and performance using value stream mapping, an agile methodology. Dell Technologies provides an unbiased, end-to-end assessment designed to deliver actionable outcomes aligned to your corporate vision and strategy, as well as a road map to get there.

- **Data-as-a-Service** offerings help create a centrally governed data catalog and lifecycle management service which deliver secure, policy-compliant access over multiple access paths to core enterprise data assets for the purposes of advanced data analytics.

- **Data Analytics Platform-as-a-Service** offerings enable data analytics teams to automatically self-provision secure, policy-compliant, sandbox environments—with their choice of tooling and data—to execute their analytic project workloads and rapidly integrate model outputs into production applications.

- **AI-enabled Framework** services include using DevOps, continuous integration/continuous delivery methods, containerization, API integrations, and open standards to rapidly operationalize analytics models into production applications.

- **Dell Technologies Deployment Services** can save you money and time by both getting your AI workflows running and maintaining their functions.

ESG found that AI leaders are 2X as likely to deeply leverage third-party services for their AI needs.[1]

Choose a Trusted Partner for AI Infrastructure

POWERFUL AND FLEXIBLE TECHNOLOGIES AND SERVICES from Dell Technologies and Intel accelerate AI workloads in data centers and across private, public, and edge clouds. Organizations that use or plan to use PowerEdge servers and Isilon storage systems powered by 2nd generation Intel® Xeon® Scalable processors can securely and reliably begin the journey to AI maturity without the need for hardware or infrastructure reconfiguration. Moreover, comprehensive data protection ensures availability and integrity of the data being used to train and feed AI models. The Dell Technologies Consulting Services team helps clients at any stage of their AI journey determine next steps for long-term success. With the breadth of the AI portfolio from Dell Technologies and Intel, you can design the best environment possible to fit your needs. With a global reach and broad ecosystem of partners, Dell Technologies can support you wherever you are geographically at any scale—from small enterprises up to the largest corporate giants.

Together, Dell Technologies and Intel can accelerate your organizational intelligence across private, public, and edge clouds with the power of AI. Visit **Dell Technologies' AI Solutions** page for more information.

Sources

1. Based on ESG Research Insights Paper commissioned by Dell EMC and Intel, "**How Organizations Unlock Their Data Capital with Artificial Intelligence**," November 2019. Results based on a survey of 750 global IT decision makers. Actual results may vary.

2. Based on ESG Research Insights Paper commissioned by Dell EMC and Intel, "**Three Transformational Compute Technologies Verified to Accelerate AI and Business Value**," November 2019. Based on a survey of 750 global IT decision makers. Actual results will vary.

3. Based on Dell EMC internal testing, March 2019, using INT8 benchmark for image classification with ResNet50 comparing Intel® Xeon® Scalable Gold 6148 vs Intel® Xeon® Scalable Gold 6248 processors on 2 socket Dell EMC PowerEdge C6420. Actual results will vary.

4. Cover image courtesy of Metamorworks/Shutterstock.