Dell Technologies and AI Dynamics MLOps Platform for Smart Factories

Revolutionizing manufacturing with AI – build, deploy and manage unlimited custom AI models easily with an intuitive end-to-end machine learning operations platform

Dell EMC PowerScale Benefits

- PowerScale enables you to innovate with unstructured data
- Scale-out architecture provides massive capacity, performance, scalability
- Seamlessly tier data across all-flash, hybrid and archive nodes
- With new NVMe drives, PowerScale delivers incredible throughput and performance for AI/ GPU workloads

AI Dynamics Benefits

- Enterprise-ready MLOps suite
- Automate ML workflows: train, retrain, a/b test and deploy
- No-code/low code machine learning architecture design for many types of ML problems
- Scale out to hundreds of containerized deployments
- Perform many tasks on a single platform – data analytics, image segmentation, image classification, video classification, text analysis, audio analysis, and many more

Solution Benefits

- Deploy on premise, in the cloud, on edge devices, or in hybrid environments
- Future-proof solution that scales to meet the most complex Smart Factory requirements
- Tight integration between Data and AI platform is must for any advanced AI solution at scale

Opportunities and Challenges for AI in Manufacturing

Factories are data-rich environments with many opportunities for utilizing artificial intelligence (AI), specifically machine learning (ML), to improve product quality, production speed, worker safety and productivity, process automation, predictive intelligence, and total manufacturing efficiency. Adopting these advanced "Smart Factory" technologies has been slow due to shortages of available AI-skilled data scientists and challenges with the scale of data needed. By joining forces, AI Dynamics and Dell Technologies can overcome these challenges by empowering manufacturers to quickly deploy smart factory technology, improving efficiency, product quality, and employee safety.

Doing more with Less

Smart factories can create new forms of efficiency and flexibility by connecting different processes, information streams, and stakeholders (frontline workers, planners, etc.) in a streamlined fashion. In fact, in a recent Deloitte survey on AI adoption in manufacturing, they found that 93 percent of companies believe that it will be a crucial technology to drive growth and innovation.

With Dell Technologies and AI Dynamics, enterprises can transform global manufacturing operations while addressing the industry's challenges of an aging workforce, growing skills gap, increased competition, and shrinking margins. Challenges associated with managing and accessing current and historical data, critical to the training of ML models effectively and efficiently, are also overcome.

By combining the power of the AI Dynamics NeoPulse platform with Dell Technologies' scalable and easy to manage PowerScale NAS, common manufacturing challenges, including maintenance and costly downtime, inventory and supply chain issues and the increasing complexity of manufacturing processes can now be addressed.

Dell Technologies and AI Dynamics for Smart Factories – Delivering a Powerful All-in-One Solution for AI in Manufacturing

Though Smart Factory and Factory 4.0, are commonly heard terms today, deployment can be a challenge for many companies. By working together, Dell and AI Dynamics solve many of the key deployment challenges, including:

Shortage of AI talent

Organizations across all industries are discovering that experienced data scientists and AI professionals are not only expensive but scarce and difficult to hire. Machine learning, and AI in general, are not for the faint at heart as they involve a steep learning curve. AI automation is a critical technology that can address this skills gap and accelerate your digital manufacturing transformation. AI Dynamics enables enterprises to create, deploy and execute AI models quickly and at scale. NeoPulse includes a built-in AUTO ML system that can tackle most real-world problems out of the box and consists of a scripting language called NeoPulse Markup Language (NML). AUTO ML was specifically created to help new users get started with just a few days of training and as little as 14 lines of code. This powerful framework can produce high cost and time savings and empower an organization to train and deploy AI models in minutes.[1]

Infrastructure complexity and interoperability

Manufacturing sites often have a large variety of machines, tools, and systems that use different technologies -- some of which may be running with outdated software or hardware that is not compatible with the rest of their environment. The absence of standards and common frameworks makes it particularly challenging for plant engineers to determine the best way to connect their machines and systems.AI Dynamics and Dell Technologies' integrated solution allows engineers to focus on delivering results, while NeoPulse manages dataset and machine learning model governance in an integrated and automated end-to-end MLOps platform. Our integrated solution supports engineers from the early stages of preparing data for machine learning to creating machine learning architectures, to training AI models, and ultimately to productizing the AI solution.

Managing factory data at petabyte scale

Manufacturing decisions often need to be acted upon immediately to identify problems before they result in unplanned downtime, safety issues or manufacturing defects. The nature of factories is such that supply chains and inputs can change on short notice.The ability to retrain ML models continuously, based on new and evolving information and historical data, is imperative. With data collection growing, data storage, management, replication, and transfer can create significant operational overhead - primarily when managed manually. AI Dynamics NeoPulse is designed to integrate directly with Dell EMC PowerScale NAS storage -- including its OneFS operating system and Data IQ and Cloud IQ features -- to deliver worry-free scalability while maintaining a historical audit trail of datasets, models, and other operational artifacts.

Edge-to-core-to-Cloud deployments

An ideal scenario for the use of AI in modern manufacturing is to be able to use data from multiple machines, processes and systems to adapt manufacturing processes in real-time. This precision monitoring and control of assets and processes utilizes vast amounts of data. This data requires ML to understand non-linear relationships to drive the optimal response to process variability. For this to be successful, low latency models need to be deployed at the edge in order to process data at the source. Furthermore, as new data becomes available, models will have to be updated automatically -- and in real-time -- via a retrain pipeline whether the data resides at the edge, core or in the cloud. AI Dynamics NeoPulse Platform and PowerScale work together to enable simplified data collection, caching, and transferring, as well as simplified model deployment to and from the network's edge.

Governance

Manufacturing companies are expected to have numerous models in production at any given time in every facility, across multiple organizational departments and geographies. The need to govern how ML models are developed, validated, deployed to, or removed from production, and how datasets and associated labels evolve is crucial to maintaining long term integrity and traceability of the entire system. NeoPulse's governance features track the provenance of datasets, models, and artifacts automatically and in the background, giving stakeholders the freedom to innovate with agility and at scale.

Trust and transparency

Without a data science background, it can be challenging to understand how data science and predictive modeling work, resulting in a lack of confidence in the abstract algorithms behind AI technology. AI Dynamics NeoPulse provides insight into the processes used to transform raw data into ingestible ML inputs. Its explainability features allow users to understand what their ML models see in the data.

---

[1] Training time is dependent on hardware speed and model complexity and can vary from minutes to hours

AI Dynamics NeoPulse:
A Powerful MLOps Suite for the Enterprise

Machine Learning Operations (MLOps) is the intersection of DevOps, data engineering, and machine learning. AI Dynamics NeoPulse combines all three elements into a single cohesive environment that can streamline the engineering, deployment, and management of AI solutions.
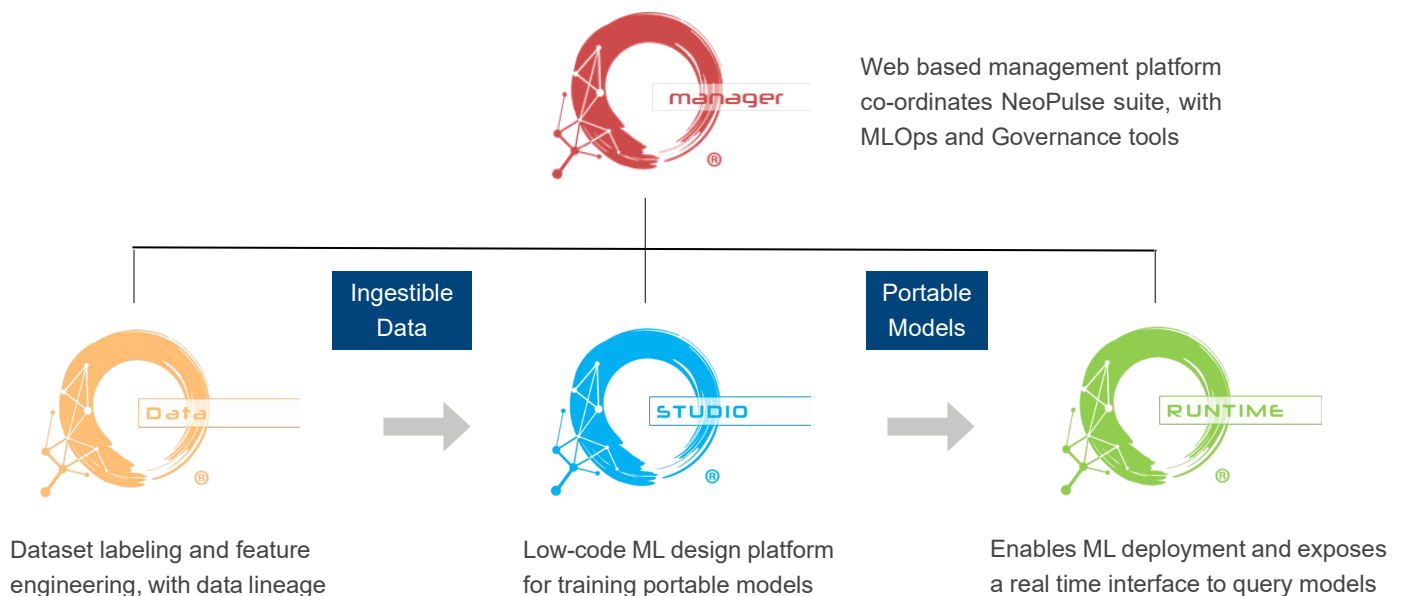
The NeoPulse Framework enables organizations to manage their entire AI workflow and infrastructure from one place. This means that DevOps, data engineers and ML engineers work from one interface instead of using separate applications. Using NeoPulse, a data engineer can assemble training data sets. The machine learning engineer can create AI models. The DevOps engineer can deploy and manage the solution without ever leaving the NeoPulse environment.
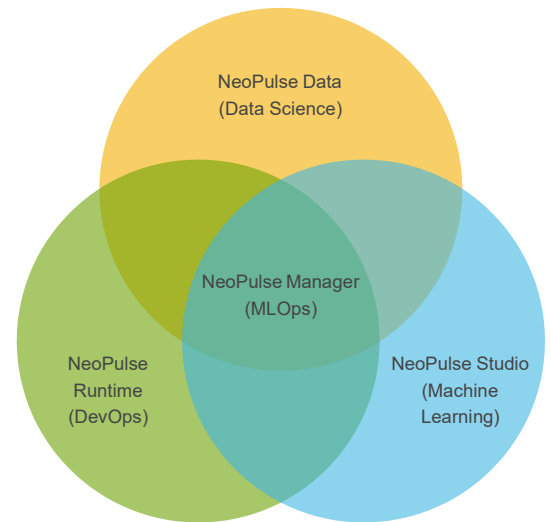
NeoPulse consists of four components that interoperate to provide complete MLOps capabilities:

- **NeoPulse Data** provides annotation tools for audio, video, text, and image data types while tracking data lineage so updated datasets can be used to train updated models automatically.

- **NeoPulse Studio** creates or imports AI models. Users can either use the onboard AutoML capability to build models automatically or hand-code the AI models using the NeoPulse Modeling Language (NML), which enables powerful model creation in as little as 15-20 lines of code. Users can also import models they have previously created using TensorFlow, PyTorch and other frameworks, and retrain them in NeoPulse.

- **NeoPulse Runtime** stages the AI model for either batch or real-time inference while abstracting away underlying hardware in the cloud, on-prem, or at the edge of ARM systems.

- **NeoPulse Manager** manages the flow of data, AI models, code and artifacts required to train, verify and deploy models. It can be used to perform A/B testing, maintains a repository of models and tracks the provenance of AI models and datasets automatically. The manager also allows organizations to deploy AI models in seconds using a web-based GUI and create models without writing a single line of code by using the graphical model builder.

NeoPulse Platform

Web based management platform co-ordinates NeoPulse suite, with MLOps and Governance tools

Dataset labeling and feature engineering, with data lineage

Low-code ML design platform for training portable models

Enables ML deployment and exposes a real time interface to query models

NeoPulse Features

Data Engineering

- Prepare training data sets
- Annotate data
- Identify anomalies
- Integrate with AI development – keep track of what data sets were used to train what model

AI Development and Engineering

- Low code/no code ML model development
- Integrate third party AI models
- Automate retraining
- Visual interface
- AI model performance baselining

AI Deployment and Management

- Organize machines, devices, data sets, applications, and models
- Drag-and-drop deployment
- Production monitoring and A/B testing
- Role-based control of assets and resources
- Deployment to the edge, on-premise, to the cloud or hybrid environments

AI Governance

- Digital Rights Management
- Provenance
- Data source integrity

Scaling

- Deploy models across hundreds of systems

NeoPulse Use Case on the Factory Floor

For one automotive parts supplier, a recall of a defective part can result in $600,000+ of losses in a single year. Moreover, a well-trained operator requires anywhere between forty-five seconds to two minutes to thoroughly inspect a part. Further compounding the issue is a decreasing pool of available workers capable of performing quality assurance tasks. All these factors mean that an AI application capable of identifying defects in milliseconds and requiring only a few seconds for an operator to decide drastically cuts the recall rate and saves approximately 58,000 person-hours for that operator leading to substantial productivity improvements.

The NeoPulse platform gives users the ability to streamline the entire machine learning process, while offering them the ability to manage data sets, control AI assets such as AI models, and automate the retraining or redeployment of AI models.

Using NeoPulse Manager, initial training uses many images of defective parts and normal parts. The images are segmented based on the location of the defect to enable the operator to identify an issue immediately. Data sets representing other modes of defect detection, such as vibration data taken from other sensors, can further refine the quality of the detection process using ensemble inference. Ensemble inference is when different AI models act in conjunction, as a committee of sorts, to increase the probability of accurate detections. During an audit if an image is falsely labeled as a defect, then the operator can submit that image for the retraining process where it will be aggregated with other data that resulted in the model making an incorrect inference.

The retrain data will automatically be sent to the NeoPulse Studio dedicated server, automatically triggering a retrain event and validation using A/B testing. If the retrained model is shown to have superior performance compared to the previous version, it will be automatically deployed to the target machine.
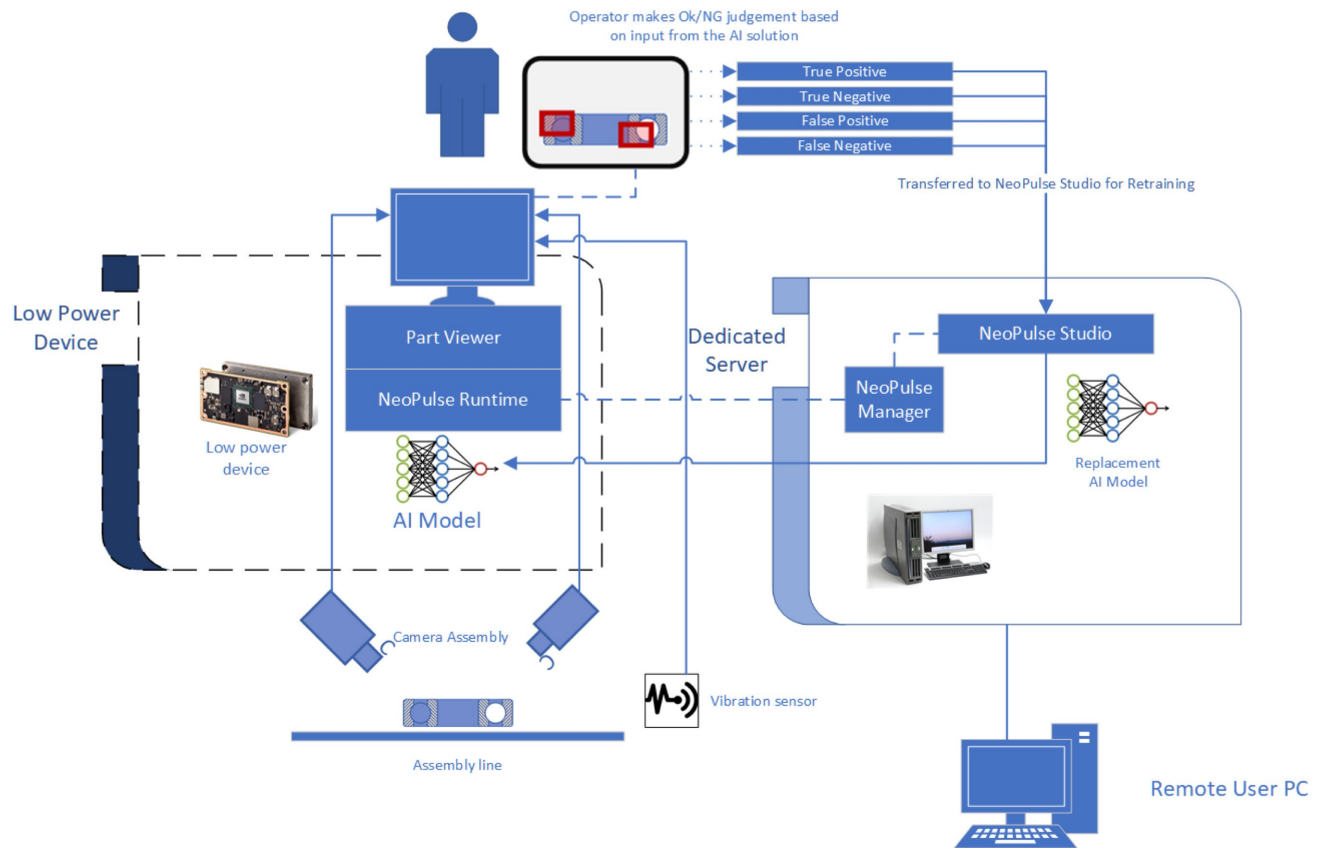
The workflow is made possible by the NeoPulse Framework, and aside from the initial training, requires very little supervision frommachine learning engineers, data scientists, or DevOps. The only manual oversight is that of the operator's periodic auditing and therefore yields substantial benefits.

Dell EMC PowerScale: The Ideal Storage Solution for AI

We see an explosion in the growth of machine learning today, which requires the collection and marshaling of vast amounts of training data that can be quickly accessed by an ML training system. Along with the accelerating growth of new data, the composition of new data is also changing significantly from traditional structured, block data to much more unstructured, file and object-based data. More than 85% of new storage capacity installed in organizations around the world is for unstructured data.

Deploying a Dell EMC PowerScale running the OneFS operating system offers organizations low latency, high availability, performant data storage, and the ability to start with a deployment as small as 11 Terabytes and scale up to 10s of Petabytes seamlessly. At the same time, OneFS manages data replication, security, and performance optimization. As more nodes are added to a system, they are automatically provisioned, so that total system performance increases at the same time as capacity.

Integrating the NeoPulse Platform with Dell Technologies OneFS can enable effective deduplication of data training sets across the organization and significantly reduce network traffic and bandwidth utilization across the company's intranet. NeoPulse on Dell PowerEdge allows enterprises to automate the process by drawing data from the edge, selecting data most likely to improve AI models, automatically triggering model retraining when new data is available, and then validating and deploying the models. If NeoPulse application deployment capabilities are used, then the entire solution deployment and management can be completely automated when certain trigger conditions are met (such as availability of new data, AI model performance thresholds being breached, etc.) using Dell Technologies PowerEdge. AI Models and applications can also be hot-swapped with zero downtime.

Furthermore, PowerScale systems can serve datasets to GPU clusters with excellent performance: a PowerScale F800 4-nodes cluster showed linear training workloads when scaling from an 8 GPU A100 system to a 32 GPU system peak throughput of over 5GB/second.
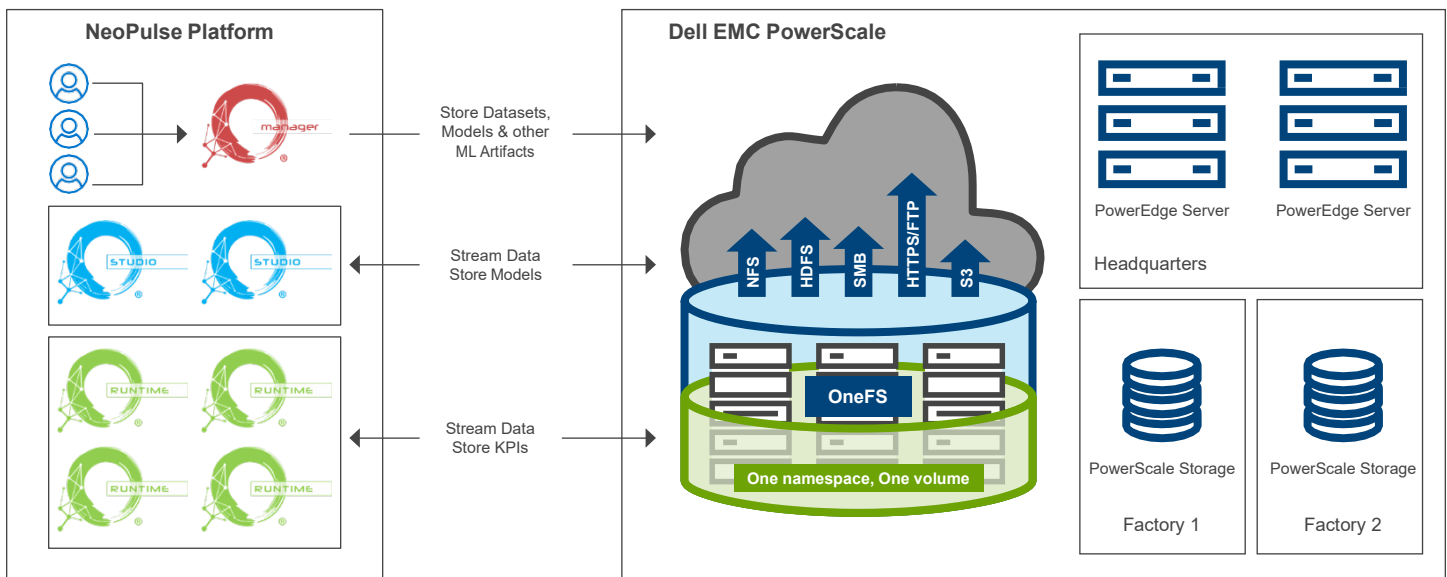
NeoPulse in a Heterogeneous Storage Environment

In a typical manufacturing setup, data may be collected at each factory and aggregated locally. This data would then need to be moved to a system that allows users to log in and add annotations. This data is then copied to the GPU system, where pre-processing and training will occur. Finally, these models can be copied back to the factory, where they will be deployed at the edge. Because we have built NeoPulse to work in any environment, including the edge, on-premises, the cloud, or an arbitrary combination thereof, data is moved between systems to ensure availability for each respective workload.

NeoPulse with Dell EMC PowerScale Unified Storage Solution

The integration of NeoPulse Manager into a Dell EMC PowerScale cluster, it changes the way in which data is handled. Because NeoPulse now knows that the storage system is both accessible by all the machines and is performant enough to serve multiple parallel workloads, including but not limited to data collection, annotation, training and inference, it can be more intelligent about how it handles the data by leveraging APIs exposed by PowerScale. Newly collected data can be passed to anomaly detection, regression, or classification models so that the system can pre-filter images that it performs the most poorly on. These images can be tagged via the OneFS object store for annotation or unsupervised training. NeoPulse Manager can do audit reports and individual customer compliance requirements leveraging PowerScale API and DataIQ API.

Because performance bottlenecks are removed, NeoPulse Manager can coordinate automatic retraining of all models that could benefit from the newly collected data by having OneFS stream this data to compute nodes. Not only does this save costs by reducing the amount of storage required at each compute node, but it significantly reduces unnecessary network utilization. Models are automatically updated when new data is ready: unlabeled data for unsupervised models and data annotated by users of the platform who are designated subject matter experts (SMEs) for supervised models. These models can then be automatically tested, compared, and deployed directly from OneFS without transferring data from one site to another. Since these new and improved models are hot-swapped, there is zero downtime for the organization, and all this can happen with no human intervention required.

Success – it's all About a Unified AI Platform

By deploying AI Dynamics NeoPulse together with PowerScale scale-out storage and PowerEdge servers, the complexity of managing datasets, models, governance, and auditability is abstracted away, allowing stakeholders to focus on delivering results that accelerate business transformation, while the underlying hardware and storage capabilities are automatically scaled up by IT staff. By leveraging this combined solution, enterprises can start out with systems as small as 11TB and seamlessly scale to dozens of Petabytes. Users will gain access to high-performance compute and scale-out storage, intelligent data storage, data mobility, multiprotocol access and highly performant streaming data access required for efficient AI model training. The combined solution results in a vastly accelerated machine learning pipeline for the enterprise, significantly lowering downtime and automating productivity improvements. The result: Enterprise-wide AI management, deployment, and governance with efficient reusable, automated model development pipelines across the organization, and ultimately, a Smart Factory.

About Dell Technologies

Dell Technologies has been a leader in the advanced computing space for over a decade, delivering proven products, solutions, and expertise. Dell Technologies has a team of data analytics, HPC, and AI experts dedicated to staying on the cutting edge, testing new technologies, and tuning solutions to your applications to help you keep pace with this constantly evolving landscape. With an extensive portfolio, years of experience, and an ecosystem of curated technology and service partners, Dell Technologies provides innovative solutions, workstations, servers, networking, storage, and services that reduce complexity and enable you to capitalize on the promise of data analytics, HPC and AI.

Click below to learn more about Dell Technologies solutions for Artificial Intelligence/Machine Learning, visit https://www.delltechnologies.com/en-us/solutions/data-analytics/machine-learning/ready-solutions-for-ai.htm

About AI Dynamics

AI Dynamics is a global organization founded on the belief that everyone should have access to the power of artificial intelligence (AI) to change the world. The company's NeoPulse® Framework is a scalable end-to-end AI platform transforming how organizations innovate and grow. Considered to be a true 'operating system for AI,' NeoPulse® is a new AI development standard that empowers anyone at any skill level to create AI solutions with minimum coding and at a fraction of the cost traditionally associated with artificial intelligence. NeoPulse® is the only AI platform that enables a standardized, reproducible process to create, distribute, and manage AI in the enterprise. A trusted solution, NeoPulse® also provides data governance and can be deployed on any on-premises device, cloud, or IoT environment.

For more information, visit https://AIDynamics.com.

References

Systems for Deep Learning: https://www.nvidia.com/content/dam/en-zz/Solutions/Data-Center/resources/dell/h18597-dell-emc-powerscale-and-nvidia-dgx-a100-systems-for-deep-learning.pdf

Reference Architecture: https://www.nvidia.com/en-us/data-center/resources/dell-emc-dgx-pod-reference-architecture/

Discover more about Dell Technologies

Learn more about Dell Technologies storage solutions for Machine Learning Applications

Learn more about Dell Technologies Data Storage Solutions

Follow Dell Technologies on Twitter

Contact a Dell Technologies Expert for Sales or Support

D⌀LL Technologies