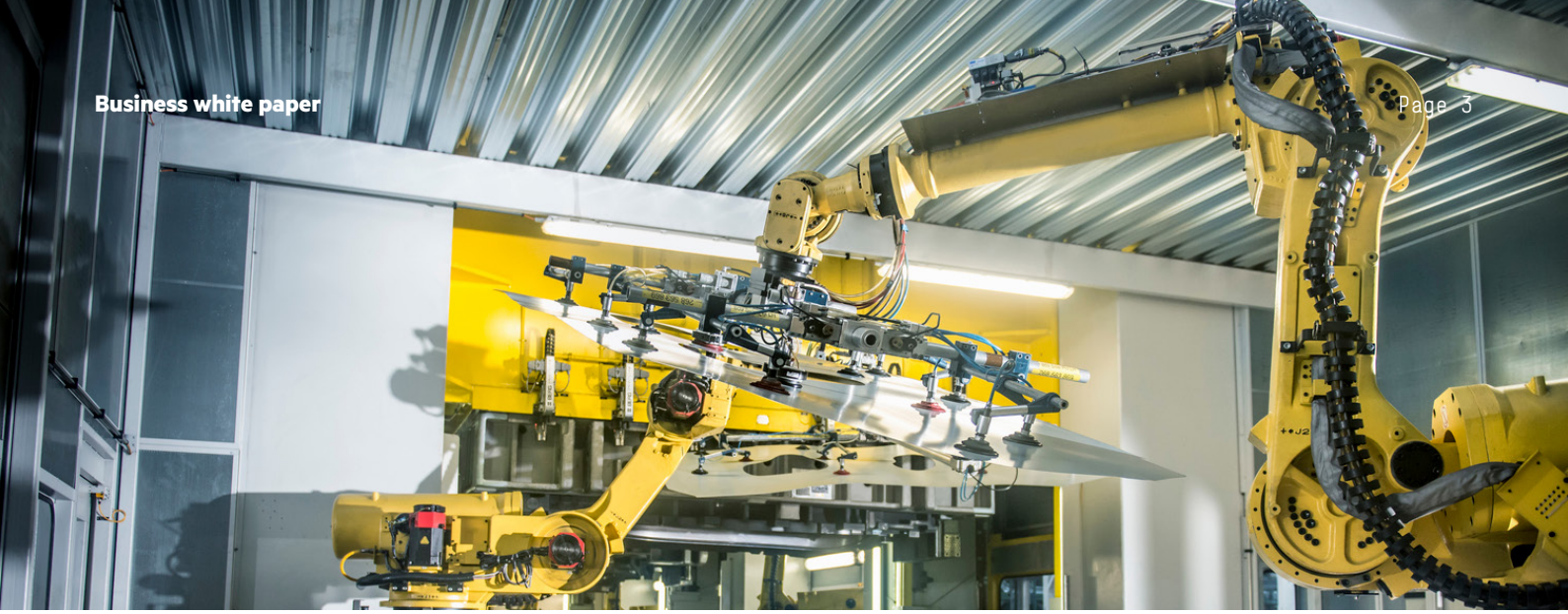
Hewlett Packard Enterprise | NVIDIA

Business white paper

POWERING THE FACTORY OF THE FUTURE TODAY

Industrial AI, where insight meets opportunity

TABLE OF CONTENTS



OVERVIEW

Artificial intelligence (AI) at the edge is the backbone of modern smart factories. Manufacturers are taking AI to the factory floor to transform their operations by harnessing insights in real time to boost efficiency, maximize profits, and move closer to zero defects.

Manufacturing is a data-rich industry and one of the first to pilot AI at the edge. Despite the vast amounts of industrial data available, research shows that roughly 90% is wasted or goes unused in manufacturing plants. This remarkable loss of information is caused by legacy equipment and operational systems that tend to be proprietary and siloed and therefore incapable of communicating with each other.

Today, the digital landscape is expanding. Ongoing technology innovation is now breaking down the barriers between data and action, bringing robust compute and processing capabilities out of data centers and down from the cloud and placing them right at the edge. Advancements in the Internet of Things (IoT), edge computing, and AI have the potential to convert unused data into tremendous value to transform manufacturing processes. With a rising number of connected devices in industrial IoT (IIoT), moving processing power from the core allows for real-time prediction and reduced network/bandwidth dependency.

As manufacturers deploy AI at the source of their data, entire production lines become intelligent. This promises to enable new business models and revenue streams. For example, AI-powered industrial robots can operate autonomously or cooperatively with humans to ramp up efficiencies and reduce costs. Connected cars are becoming smartphones on wheels, causing traditional automakers to reposition themselves as mobility and technology providers to remain competitive. Using AI algorithms, companies enable their workforces to boost yields and manufacturing capacity by developing new processes and service-oriented business models.

The new age of manufacturing will see a major paradigm shift that connects edge to cloud, fueling insight and innovation on an unprecedented scale. Companies at the forefront of this movement are taking AI to where it matters most.

WHY AI AND MANUFACTURING ARE SUCH GOOD FITS

Manufacturing is entering a new phase of production centered on customization and process automation. Companies strive to operate with increased agility and responsiveness in order to innovate quickly and tap into new sources of revenue. However, there are significant challenges standing in their way.

Exponential data growth is creating a choke point for manufacturers. It takes robust compute performance to aggregate and process data, and the infrastructure is traditionally housed in a central location, away from the operational technology (OT) used on the factory floor. Sending volumes of data back to the data center or cloud for processing can delay time to insight and leave factories waiting for time-sensitive inputs, which has a direct negative impact on operational efficiency.

New advances in data processing, real-time analytics, and other IT technologies have prompted innovative OT approaches that provide insights when and where they are needed most. The convergence of IT and OT is an opportunity for manufacturers to bring rapid insights and competitive advantages to the industrial edge. Manufacturers recognize that the success of their products in the market is based on their ability to collaborate and share inputs seamlessly across a network. Companies must be proactive to bridge these two worlds and unlock substantial productivity gains. At the same time, intense competition and escalating demands for new high-quality products have caused more urgency to build an operating environment that converges IT/OT and is faster, more intelligent, and highly adaptable.

Manufacturers are implementing AI and analytics to meet these challenges of production. With the latest technologies and tools providing increased agility and deeper insight, companies have the power to enhance industrial processes, maximize automation, and achieve better results at the edge.

AI AT THE MANUFACTURING EDGE

The manufacturing industry has been at the forefront of automation with Industry 4.0, and companies with edge-centric and data-enabled strategies will continue to lead the pack. For Industry 4.0 to reach its full potential, manufacturers need to take advantage of the enormous amounts of data generated by countless edge devices. Leveraging AI in the factory is a big step toward making entire operations smarter. By integrating historical performance data with real-time analytics, AI makes it possible to predict how products will perform in the real world, when plant equipment will need maintenance before breakdowns occur, or to catch product defects faster and more accurately than the human eye.

AI has traditionally been confined to the data center, where powerful computers run complex algorithms and send results back across the network. Today's manufacturers are decentralizing IT and investing in breakthrough techniques such as machine learning (ML) and deep learning (DL), which are rapidly making their way to devices along the assembly line. AI at the edge can help manufacturers make real-time decisions in a matter of milliseconds and automate key processes to improve production performance.

Manufacturers use a metric called Overall Equipment Effectiveness (OEE) to gain visibility into the variables in this chain: machine availability, performance, and quality of output. It is often referred to as the "gold standard" for measuring manufacturing productivity. Smart manufacturers are now improving OEE in production environments through greater visibility and control. AI captures streaming data from smart equipment, cameras, and sensors on the assembly line to quickly extract insights that direct factory operations, maintenance, and quality control activities. Using real-time datasets, AI can identify issues and recommend real fixes to help manufacturers produce high-quality parts with little to no downtime.

Without AI algorithms to monitor production, machines continue to run even when they encounter maintenance or operational issues. This lack of intelligence puts them at risk of becoming damaged. Smart machines can make decisions about whether to stop when insights are generated continuously, rather than going back to a data center and waiting for an operator to give instructions. Making that decision immediately and even autonomously at the machine helps manufacturers prevent unnecessary loss and downtime and operate with more efficiency. This type of intelligence is necessary to reach an OEE score of 100%. The graphic below is a view of a smart factory which shows how IoT and AI drive increased availability, quality, and performance.

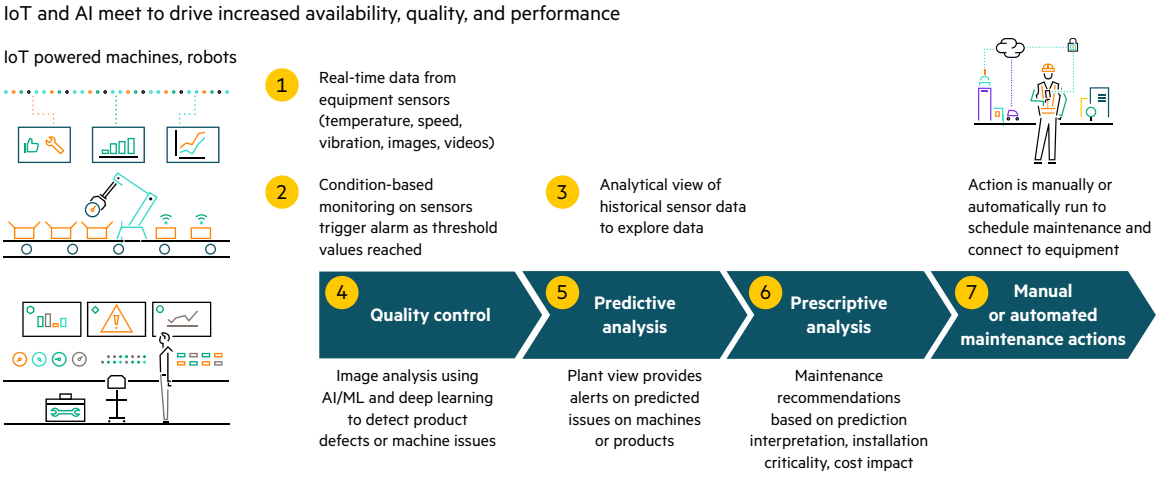


FIGURE 1. A day on a factory floor

AI USE CASES IN MANUFACTURING

AI powers the most impactful use cases in manufacturing today. These include video analytics, predictive and prescriptive maintenance, and quality assurance (QA). Moving forward, the industrial use of AI will become an increasingly important tool for enhancing industry use cases and transforming how companies work, collaborate, and innovate.

Video analytics

Video analytics powered by AI continues to expand within manufacturing as the industry looks for new and more effective ways to improve production and drive profit margins. The use of ML and DL to analyze live video feeds delivers greater visibility and real-time insights at each stage of production, enabling companies to make adjustments and resolve issues before they impact products or customers. These techniques target applications that rely on computer vision technology—such as predictive maintenance, quality management, worker safety, and autonomous operations utilized in smart factories.

Creating an effective video analytics solution begins by accessing the right data, properly curating it, and converting and annotating it. Manufacturers depend on AI to integrate insights from historical data with data streaming from the edge to gain a complete picture of their operations. AI algorithms leverage the information to label images and subjects and become increasingly effective at recognizing when specific events occur. Performing these tasks at the source of data accelerates time to solution and increases the accuracy of results. This enables manufacturers to capture important insights from video and reduce false positives.

Building trust in data and training AI models with a feedback loop are both critical to this process, which helps deep neural networks learn from their successes and mistakes. AI activities can be coordinated across many edge locations with the use of centralized fleet management. Centralized control planes provide a single view of data, where manufacturers can perform model training and video analysis. Input for the centralized AI is collected from multiple locations, and the output is then broadcasted back to the edge.

To overcome manufacturing challenges, it is important to develop an approach to AI that accelerates video analytics workflows. When building such a solution, you should consider cameras, sensors, controllers, and management systems, along with edge systems, data center assets, and enterprise storage. Appropriate network infrastructure is required as well. Top manufacturers are developing advanced deep neural networks built on high computational power and the massive acceleration of GPUs right at the source of their data to process complex video files. When combined, these powerful technologies make it possible to acquire and aggregate images from multiple cameras for analysis, and rapidly train AI models at the network edge to streamline decision-making and timely actions. Now, machines are able to learn at the extreme speed, accuracy, and scale that enable true AI.

Predictive and prescriptive maintenance

AI plays a vital role in equipment maintenance. The ability to process data where it is created allows systems on the factory floor to identify patterns in equipment condition and recognize anomalies almost instantly. AI leverages these inputs to predict potential maintenance issues in real time and alerts manufacturers before they have an adverse effect on production. The systems deliver specialized recommendations and corresponding outcomes that reduce operational risks and improve OEE. In smart factories, AI is now used to take the next step to autonomously resolve problems without the need for human intervention. These capabilities are empowering manufacturing environments with increased uptime, productivity, production efficiency, and safety.

Predictive maintenance is a core activity of this process. By collecting precise sensor data describing an asset's condition and overall operational state, AI can compare current data with observations from the past—such as machine or component failures and nonfailures—to predict future events. It can even make and implement recommendations for maintenance that prevent events from occurring. These insights significantly cut back on required maintenance and costs. Today, AI/ML deployments along the assembly line can analyze a vast amount of data in real time to predict failures and manufacturing discrepancies.

While predictive analytics tells manufacturers why equipment is likely to fail, prescriptive maintenance uses AI to determine the best action and outcome. Prescriptive maintenance enabled by AI/ML models can predict and prescribe upcoming maintenance cycles and automate the scheduling of maintenance. This capability is paramount for improving asset availability, performance, and quality associated with OEE. For example, using AI techniques, a system can use operational and historical data to predict that an asset is going to fail with concrete failure time frames. Based on contextual business data, AI can also analyze the impact of failure and the cost of a proactive fix. This not only increases the accuracy of predictive maintenance outcomes, but it also guides strategic decision-making and provides suggestions about actions to consider. Manufacturers can then link the prediction manually or automatically to a predetermined action, generating a maintenance ticket or even shutting down the equipment.

Quality assurance

AI has enabled dramatic advances in QA by replacing manual inspection with data-driven inspection processes that are faster and more precise. Companies have an opportunity to match the speed of automated manufacturing with automated quality control and management. By training AI algorithms based on a highly precise set of standards, manufacturers can detect process deviations and anomalies early in the process. Manufacturers can make ongoing adjustments and optimizations to increase first-pass yields while lowering the overall number of defects. These groundbreaking capabilities are critical to increase QA and accelerate the development of high-quality products. The benefits are undeniable:

- **Increased yield:** Boost efficiency and production throughput while reducing scrap
- **Consistent quality:** Enable repeated high accuracy under varying conditions
- **Fewer false negatives:** Optimize production costs and minimize downstream waste
- **Reduce false positives:** Decrease the need for manual reinspection with improved test sensitivity

Today, AI/ML algorithms provide an extreme level of accuracy when identifying product defects on images and video. With this methodology, the entire task of quality inspection can be automated right at the edge, and fault detection can be accomplished instantly.

REAL-LIFE EXAMPLES OF AI IN MANUFACTURING

Manufacturers around the globe are adopting the latest technologies and tools to address specific AI use cases. Below are three examples that are helping companies solve their greatest challenges.

Moving closer to zero defects

Relimetrics is helping manufacturers enhance how they design and create products with software that fully digitizes their quality audits. The objective is to enable companies to achieve zero defects during production. To make this possible, Relimetrics needed an AI solution that could automate their customers' quality management processes and deliver critical insights on demand to improve product development.

Relimetrics implemented computer vision technology to capture and analyze data on the shop floor in real time. With a combination of video analytics and ML, they harness insights at the edge that help to reduce errors, recoup investments, and maximize product output. This is key to keeping up with demands for rapid production and product customization, increasing customer satisfaction, and remaining competitive in fast-paced markets.

An edge-to-cloud architecture was engineered to save time and prevent latency problems by performing inference at the network edge. Relimetrics selected the HPE Edgeline Converged Edge System equipped with NVIDIA® GPUs to deliver massive acceleration for ML and analytics workloads. This is particularly important because each camera installed at the conveyor belt streams large volumes of data (about 3 GB per hour). The solution includes NVIDIA TensorRT software for high-performance ML inference along with the NVIDIA Metropolis framework, which simplifies the development, deployment, and scale of AI-enabled video analytics applications. These technologies have an unmatched capacity to convert data from trillions of AI and IoT devices on factory floors into valuable insights.

To save time in training ML algorithms to detect issues or anomalies, Relimetrics and HPE Pointnext Services experts built the machine vision system at the Foxconn facility to require the storage of reference images of server components instead of whole units. With this design, the manufacturing execution system provides a bill of materials for each product on the conveyor belt so that the system can make a full reference image based on reference image components.

Relimetrics decreased inference latency from 4,165 milliseconds down to 3—a performance improvement of nearly 1400x. Now, manufacturers can inspect the configurations and properties of their product components with extreme speed and accuracy.

Building the refinery of the future

Texmark Chemicals, Inc., is a crucial link in the petroleum product supply chain, specializing in contract manufacturing as well as the production of specialty and high-volume chemicals for the world's leading chemical companies. The company wanted to find better methods to drive plant performance and drive down costs by enhancing OEE. However, limited visibility made maintaining a safe and effective operating environment costly and time-consuming.

Texmark needed a technology infrastructure capable of producing the speed and processing power to run analytics at the edge. The ideal solution would enable them to monitor the condition of connected assets and predict when failures would occur to reduce unplanned downtime, extend equipment life, and optimize machine performance.

To meet these challenges, Texmark needed a high-powered AI solution to convert data into actionable insights and allow workers in different locations to quickly operationalize results. They launched a multi-phase project to deploy a comprehensive IIoT solution, combining the compute performance of the HPE Edgeline Converged Edge System with an NVIDIA T4 GPU for predictive maintenance, asset integrity management, and facility condition monitoring. The result was an asset management solution purpose-built for training AI models for a wide range of applications—including risk-based inspection activities, anomaly investigation, and resolution procedures.

By rapidly collecting and analyzing process data from censored devices, the company gained visibility across its operating environments, streaming insights from the edge to determine when and where to perform preventative maintenance. Leveraging AI-powered insights enabled Texmark to achieve smart plant-level performance and reduce planned maintenance costs by 50%. The company continues to improve their bottom line through maintenance savings, uninterrupted productivity, and increased production uptime.

Automating quality inspection processes

Seagate Technology is the world's largest manufacturer of hard drives and a leader in ensuring the quality and reliability of their products. They specialize in high-value manufacturing products using high-cost capital equipment. This means Seagate's factories produce huge data volumes—approximately 5 TB per day—from millions of images and thousands of factory tools sensors. With large, complex data sets streaming from numerous locations, the company struggled to perform quality inspections due to high latencies and the limitations of rules-based analytics. To maintain the highest quality standards, Seagate needed to eliminate the wait times and bandwidth costs of using cloud analytics in order to increase quality inspection throughput and scale globally. This called for a new solution that could support smarter, faster analytics at the point of production.

Seagate's goal was to improve manufacturing QA using AI at the edge to find defects early. Their engineers implemented an AI solution that eliminates inefficiencies and prevents anomalies in products before they are made. AI automates the collection of equipment, operational, and inspection data in highly complex manufacturing processes with many stages, at multiple sites. As the technology becomes incorporated, Seagate expects to see greater efficiency and quality, enabling them to address anomalies in real time and minimize costly false positives.

One solution is powered by the HPE Apollo 6500 Gen10 server that is optimized for DL training in a wide range of AI applications. It includes eight NVIDIA GPUs to accelerate neural network models and deliver up to 125 TFLOPS of single precision compute performance. Another solution utilizes the HPE Edgeline Converged Edge System, designed for large-scale inference workloads for factory floor deployments including IIoT, edge computing, and AI inference apps with up to four blade servers. Each blade has one NVIDIA GPU capable of running multiple inference models simultaneously. HPE Pointnext Services developed a Kubernetes-based infrastructure that supplies Seagate with the flexibility to scale their models and manage them seamlessly.

With this solution, Seagate is taking QA to new heights. In one project at a factory in Thailand, engineers estimate a 20% reduction in cleanroom investments, a 10% reduction in manufacturing throughput time, and up to a 300% return on investment (ROI) from improved efficiency and better quality.

HOW TO GET STARTED WITH AI

With roughly 70% of apps and data still residing on-premises, the HPE GreenLake platform delivers the cloud experience across manufacturing edges, collocations, and data centers. Pioneering systems, apps, and models from NVIDIA (ngc.nvidia.com), combined with the expertise and comprehensive set of computing and infrastructure breakthroughs from Hewlett Packard Enterprise (HPE) can help you unlock the value of AI/ML, and HPC.

With NVIDIA Metropolis and HPE, companies can bring powerful AI-based video analytics applications for factories to market faster. Developers of vision AI applications face significant challenges, including model training and optimization, pipeline throughput and managing apps/fleets of compute nodes at the edge. NVIDIA Metropolis is an end-to-end application framework that brings visual data, edge computing and multi-modal AI together for developers creating AI solutions that improve operational efficiency and safety for factories.

Through deep collaboration, HPE and NVIDIA have developed powerful, intelligent, and data-driven solutions to help companies revolutionize their production environments and scale rapidly to meet industry demands and prepare for what comes next. Our winning combination of technologies enables staggering compute power and speed for the most common use cases in manufacturing. They are proven to enhance AI applications and accelerate valuable insight in any industrial environment. Together, HPE and NVIDIA are making the factory of the future a reality today.

Once manufacturers have selected the best-fit HPE and NVIDIA technologies for their requirements, we have an incremental three-phased approach to ensure success. Regardless of your starting point, this strategy is proven to support AI adoption. We make it easy to get started with AI using these simple steps:

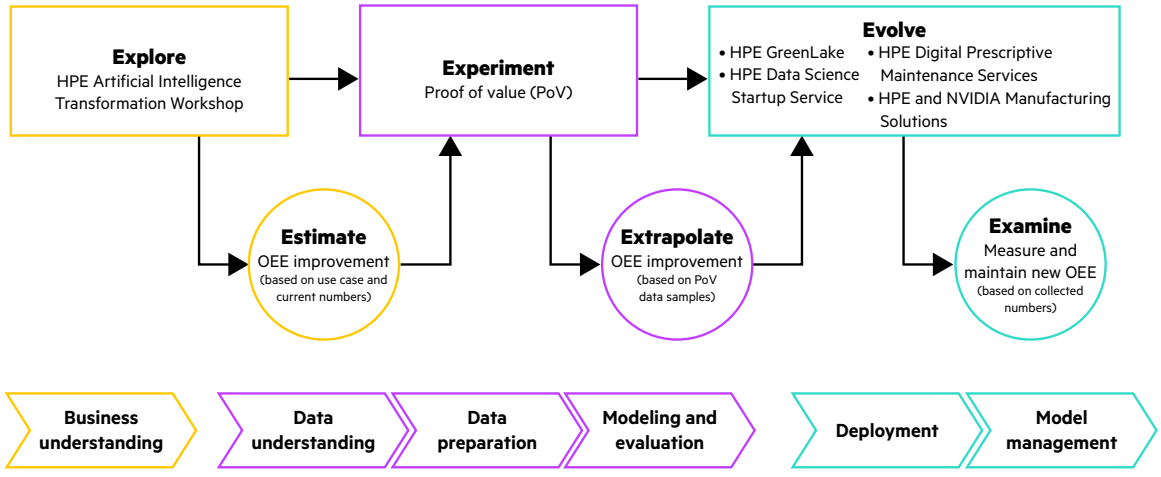


FIGURE 2. An incremental approach to AI

Explore

The first step is to identify the opportunities AI will bring to your production environments. Companies need to determine the AI use cases that are essential to their manufacturing processes and ensure they have access to the right data to achieve the results they need. HPE Artificial Intelligence Transformation Workshop are available to prepare your teams, explore AI including data opportunities and success factors, and create high-level customized plans with recommended use cases for testing. HPE experts can guide you through one- or two-day workshops to align business, data, and IT stakeholders.

Experiment

In this phase, manufacturers begin building a proof of value. This means proving that AI will add value to your operations by applying models to a significant amount of data. Experimenting determines how you can gain intelligence from historical, operational, and business data and how to effectively integrate these insights into existing manufacturing processes. Adopting the appropriate tools and techniques is vital to the success of your AI applications and must be integrated into your existing systems.

HPE Executive Briefing Centers (EBCs)—HPE EBCs offer select manufacturing customers access to cutting-edge technologies and expertise, including HPE GreenLake and solutions based on the latest NVIDIA GPUs on HPE systems. IoT fast-start solutions and AI proof-of-value services are also available to support your experiments. HPE Pointnext Services has developed pre-integrated production pilots that deliver immediate insights into critical assets. You can test use cases such as video analytics, predictive/prescriptive maintenance, and quality control and generate results without committing much time or resources. These pilots include consulting services, HPE, NVIDIA, and strategic partner technologies to help you adopt AI and you decide to scale up or scale out.

Evolve

Once you have defined a project goal and tested your results, it is time to evolve. In this phase, manufacturers deploy the first AI model into production with the right deployment model, learn to maintain model quality over time, scale and consume the solution on-premises with more use cases. Most AI pilots fail to scale due to data related challenges. It is critical to address these challenges to unlock the value of data in IT and OT systems.

HPE GreenLake brings the cloud experience to your apps and data across edges, clouds, collocations, and data centers—for you to discover, access, store, prepare, manage, protect, analyze, and act on your data across the entire lifecycle. HPE GreenLake for HPC offers superior flexibility and control to power AI use cases at an unprecedented scale. With HPE GreenLake, organizations can choose an on-premises HPC solution running on NVIDIA GPUs with all the flexibility, scalability, and utility-like consumption of the cloud. Customers can increase their agility with pay-per-use pricing and pre-installed buffer capacity that is ready to provision when needs grow. Whether demand for HPC resources spikes suddenly or grows steadily, manufacturers are always ready to meet their new needs.

HPE GreenLake for ML Ops solution brings DevOps agility to every stage of the ML lifecycle, and frees up your data scientists to concentrate on data science by easily accessing open-source or third-party independent software vendor (ISV) tools as well as required data sources across their organization. The solution runs a variety of NVIDIA GPUs including A100 for AI/ML training and inferencing use cases in both HPE Apollo and HPE ProLiant servers, which are all NVIDIA-certified. Included with HPE GreenLake ML Ops, HPE Ezmeral software simplifies the development, scaling, and management of Kubernetes clusters and machine learning models across hybrid environments. HPE Ezmeral is an open data and analytics software that brings cloud agility to AI workflows at the edge. The ability to run AI applications using open-source Kubernetes—analyzing and processing data natively—lets you spin up new instances with the press of a button.

As you expand the use of AI, NVIDIA Omniverse™ keeps you connected and productive. Omniverse is a cloud-native platform built for collaboration and real-time photorealistic simulation. Geographically distributed project teams can bring together data from multiple sources for numerous AI applications—from ideation and conceptual design to production automation on the factory floor. Manufacturing in Omniverse connects everyone and everything throughout the process. With this platform, manufacturers can develop connected virtual factories that are innovative and intelligent.

GETTING STARTED

HPE technology stacks and blueprints for manufacturing built on leading AI, HPC, and edge computing, NVIDIA technologies, and HPE Pointnext Services and a rich ecosystem of software and solution partners enable you to accelerate deployments edge to core.

HPE Pointnext Services brings together a full suite of capabilities—from advisory to implementation and operations—to accelerate AI adoption at the industrial edge. Analytics workloads happen at the edge for fast action, and often this data needs to be integrated into the existing data environment or a new data platform environment. This is where HPE experts can help you explore, experiment, and evolve your AI solutions for manufacturing use cases. We work with you to define and validate use cases, connect and protect edge IoT platforms, integrate and analyze manufacturing data, and apply AI techniques where models can automate key processes, predict and prescribe upcoming maintenance cycles, and streamline product development.

CONCLUSION

HPE and NVIDIA are bringing breakthrough intelligence to every production environment, from the supply chain to the factory floor. We believe success requires constant evolution. With an extensive portfolio of technologies and support services built on years of collaboration and deep industry expertise, we empower companies to evolve with confidence and prepare for tomorrow with the capacity to compete and lead.

We can help you pioneer the new age of manufacturing. Discover how you can utilize AI at the edge, where insight meets opportunity.

Make the right purchase decision. Contact our presales specialists.

Chat now (sales)

Call now

Get updates

LEARN MORE AT

hpe.com/us/en/solutions/artificial-intelligence

hpe.com/us/en/solutions/manufacturing

hpe.com/us/en/solutions/hpc-high-performance-computing/nvidia-collaboration

Hewlett Packard Enterprise

a50005170ENW, Rev. 4