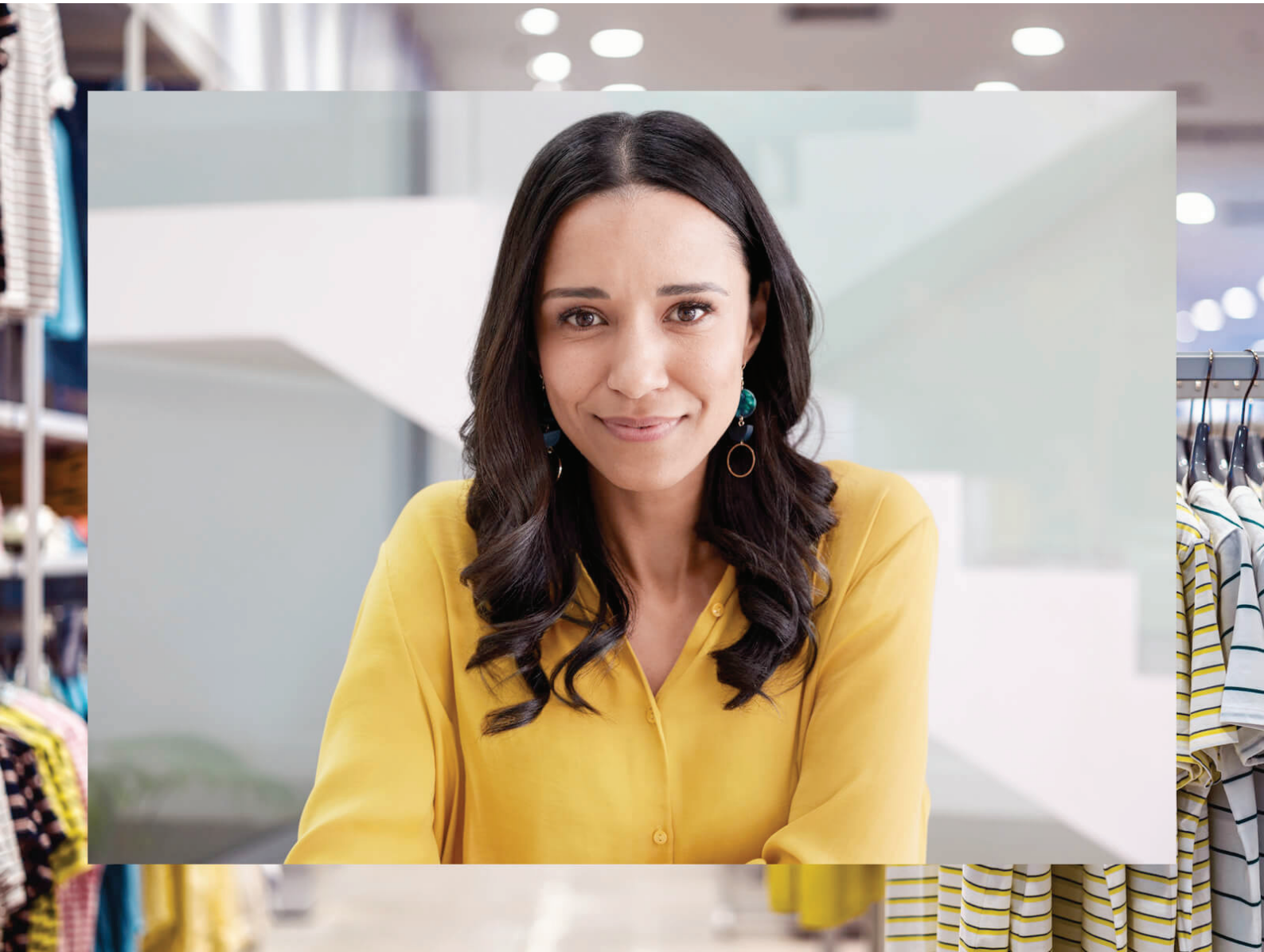
Hewlett Packard Enterprise | NVIDIA

Solution overview

Empowering a new era of smarter retail

Accelerate retail operations with AI from edge to cloud



Retailers are using the power of AI to transform how they work and deliver compelling shopping experiences:

- Anticipate consumer needs.
- Accelerate insights at the near edge.
- Mitigate shrinkage and loss.
- Enable store analytics for real-time decisions.
- Automate and optimize the supply chain.
- Decrease labor and cost challenges.
- Improve the overall consumer experience.
- Ensure employee and consumer safety.

Uncovering opportunities in retail

The retail industry is in a state of constant change as companies flood the market with new products while trying to manage shifting behaviors, escalating consumer demands, shrinking margins, and huge amounts of data. But these challenges also bring important opportunities.

Many retailers are implementing new technologies and business models to improve brand loyalty and increase profits. Today's consumers are highly sophisticated and expect shopping experiences that are frictionless and personalized, allowing them to quickly locate and purchase the products and services they need. Providing omnichannel convenience without compromising security is critical to convert more sales in store, online, and on mobile. Boosting profits doesn't stop here.

Mitigating shrinkage and loss is an industry-wide priority—problems which are caused by theft, transactional errors, fraud, and product damage. Retailers worldwide are losing over $100 billion per year due to shrinkage, and half of that is happening in North America. Retailers must find ways to operate more effectively to create better consumer and employee experiences.

Savvy retailers are driving business agility and expanding profit margins with a breed of artificial intelligence (AI) built specifically for them. Operations powered by AI achieve better-than-ever efficiency across the supply chain and capture timely insights at the edge, where seconds count.

Driving business value with AI

By tapping into AI, retailers can power intelligent stores, omnichannel management, and intelligent supply chains. AI-enabled environments offer key advantages:

- Enhancing end-to-end operations with real-time speed, predictability, and accuracy
- Utilizing data in the store, where it is created
- Leveraging faster intelligence to solve multiple problems in a unified model
- Securing regionalized retail IT at scale and under budget

These capabilities can increase annual revenue by more than $1 trillion. Retailers are unlocking this value with critical AI use cases.

Applying AI in retail operations

Intelligent stores utilize AI to gain deeper visibility into store operations and consumer behaviors. Data generated from point-of-sale transactions, cameras, and sensors offer rich insights to enhance the shopping experience—for customized shopping experiences, merchandise velocity trend analysis, and real-time employee alerts for misplaced or damaged items. AI leverages this data to determine consumer needs and buying preferences, deliver revenue-driving online and in-aisle promotions, and provide autonomous shopping options like mobile checkout, AI-enabled carts, nano stores, smart cabinets, and fully autonomous stores.

Fraud detection is helping to protect retail profits. Applications such as video analytics accurately and efficiently increase asset protection and reduce shrinkage by preventing issues like ticket switching, mis-scanning, return fraud, and employee theft. With increased visibility at points of sale, retailers can reduce shoplifting and other types of loss in real time.

HPE and NVIDIA enable retailers to gain deep insights faster to maintain a competitive edge: Our pioneering systems, software, and pre-trained models, combined with the expertise and comprehensive set of computing and infrastructure breakthroughs from HPE, can help retailers streamline AI workloads across the supply chain to accelerate to data-driven transformation.

Conversational AI can improve consumer satisfaction and achieve better customer service outcomes. Applications such as virtual assistants, digital avatars, and chatbots are bridging the gap between humans and computers, using natural language processing and machine learning (ML) for language understanding and speech analysis. In call centers, AI-enabled chatbots are used to answer frequently asked questions and automate routine tasks, so agents can focus on consumers that require hands-on assistance.

More retailers are utilizing recommendation systems and visual search features to drive more sales. Intelligent recommendations and augmented reality create highly tailored experiences, offering features like personalized promotions to deliver live shopping suggestions that can expand cart sizes with opportunities for upselling and cross-selling. AI recommenders across different channels account for as much as 30% of revenue, which translates into billions of dollars in sales.

To support an omnichannel experience, many online retailers are deploying their own account/payment systems. Online retail is a vast and fast-growing marketplace that demands seamless shopping and convenient checkout methods. Using traditional banks for account management and payments can limit the options provided to online consumers. Moving to the cloud enables payment platforms to accommodate business growth while remaining compliant with industry payment standards.

Supply chain optimization can reduce operational costs by up to 60%. Retailers use a combination of AI, predictive analytics, video analytics, and robotics to ensure that quality products are delivered to consumers as efficiently as possible. Downtime can cost $3,000-$5,000 per minute. Greater insight and visibility streamline the product journey from demand forecasting, inventory management, and warehouse logistics to routing optimization.

Introducing purpose-built AI solutions for retail

Hewlett Packard Enterprise and NVIDIA® are empowering a new era of smarter retail with the latest advances in AI. Our solutions are more than just products. We deliver an industry-leading portfolio of optimized AI solutions that provides the performance, security, and flexibility to transform retail operations from edge to cloud. Together, these offerings create a complete intelligent platform that turns data into immediate insights and competitive advantage.

The AI platform for retail from HPE and NVIDIA combines proven and emerging AI solutions, including all the tools and capabilities to compete and thrive in the future of retail. Each component is carefully chosen to make the most of AI's potential, so companies can innovate with confidence.

At the foundation of this stack is a broad portfolio of HPE ProLiant Gen11 systems. We bring together exceptional compute powered by the latest NVIDIA GPUs in validated configurations for speed, scale, and manageability. The HPE ProLiant Gen11 portfolio is designed to optimize IT investments with compute capabilities built for companies at the edge. Combining a cloud-operating experience and built-in security, retailers can operate smoothly with the capacity to run GPU-accelerated applications and deliver unique technology solutions, develop new business models, and increase operational performance. These platforms are optimized for diverse AI workloads and backed by enterprise-grade support from HPE and NVIDIA that makes us trusted partners in the retail industry.

HPE Ezmeral Data Fabric is the backbone of the AI platform. Data fabric technology helps to monetize data wherever it exists. With seamless access to data scattered across hybrid environments, retailers can easily process information and fuel decision velocity in any location. Robust security and real-time identification provide secure access to analytics-ready data sets for developing AI models that retailers can trust. Multiple apps and users can reuse the same data sets, which reduces the amount of infrastructure dedicated to analytics. With broad support for industry standard APIs, retailers can simplify data management while gaining more value from existing data. The fabric also supports a variety of opensource tools for analyzing data wherever it lives—across on-premises, multiple clouds, and edge environments.

NVIDIA AI Enterprise is the framework that accelerates AI workflows at scale. The suite features end-to-end AI software that enables rapid model training and inference in edge and near-edge environments. NVIDIA offers an extensive catalog of cloud-native software to ramp up the building and deployment of AI models. The library includes AI solution workflows, frameworks, pre-trained models, and infrastructure optimization tools for putting AI into production anywhere. With these offerings, retailers can remove artificial silos by creating a shared pool of resources and enable on-demand, self-service GPU-accelerated analytics environments to jump-start the use of AI and start gaining ROI quickly.

HPE is integrating Pachyderm with our existing solutions to expand AI-at-scale capabilities. Running on the data fabric, Pachyderm software streamlines data processing and workflow management and enables reproducible AI. Pachyderm provides historic data insights while only processing today's new data to save time and energy. For retailers, this means unleashing large-scale AI applications and automating ML pipelines to accelerate the entire AI journey, from POC to production.

A key enabler of the AI framework, HPE ML Development Environment (MLDE) offers the fastest way to create enterprise-scale ML models. Retailers can automate model training, sharing information quickly and securely across disparate teams using AI compute clusters. HPE MLDE reduces the complexity and cost associated with model development and optimization, so retailers can deploy more accurate models, manage GPU costs, and reproduce experiments to ensure the best possible results.

Data centers need high-performance storage solutions to handle endless streams of data in an efficient way. HPE integrates storage solutions for AI that allow full performance and capacity at scale with cost-effectiveness built in. By simplifying data management and providing immediate access when it matters most, retailers can keep moving forward, faster.

To maximize productivity, HPE and Aruba networking technologies create mobile experiences at the edge with fast, stable connectivity. High-speed networking with built-in AI features offers unparalleled intelligence and automated troubleshooting to help optimize and secure the network. Retail environments benefit from increased manageability and security, so employees can work seamlessly and securely across stores, warehouses, distribution centers, and in between.

Retailers can adopt the AI-platform-as-a-cloud model to keep pace with the demands of AI while reducing infrastructure costs. HPE GreenLake offers the agility and flexibility of the cloud with the security, simplicity, and control of on-premises IT. Retailers can scale on a pay-per-use basis using a variety of services to extend AI capabilities from edge to cloud:

- HPE GreenLake for payments supports card payments and alternative payment types, such as contactless and buy now pay later, with low risk and high performance and security at scale.

- HPE GreenLake for AI and analytics can streamline and de-risk AI projects with the ability to automate monitoring, scale accordingly, and orchestrate remediation where necessary.
- HPE GreenLake for ML Ops makes it easier and faster to deploy ML workloads on-premises with an open, GPU-based infrastructure and cloud experience.
- HPE GreenLake for ML offers an enterprise-grade ML cloud service to rapidly build, train, and deploy ML models—from pilot to production, at any scale.
- HPE GreenLake management services monitor, operate, and optimize infrastructure and applications, delivered globally to provide unified control.

HPE services are available to accelerate every stage of the AI journey, from planning and deployment to leveraging AI-at-scale for the fastest-growing workloads. HPE experts help to define business goals and identify roadblocks, so retailers can implement the right innovation strategy and solution for their needs. As part of the AI strategy, HPE delivers a broad set of tools for enabling data prep and management. Unified data components support data, metadata, access, and orchestration. ML components are available to support the entire AI pipeline, from provisioning all the way to production.

Partnering with the industry's top ISVs

HPE and NVIDIA work with a number of independent software vendors (ISVs) to assist in launching AI strategies, products, and services. Our rich ecosystem offers proven tools and applications to tackle the most demanding retail challenges.

Shrinkage is an ever-evolving problem that can be mitigated with the right technology—the trick is to maintain the consumer experience and drive even faster analytics. Everseen offers visual AI solutions to transform how retailers see and solve problems. Their pre-built solutions can be deployed in front of store, back of store, and on the shop floor to enable loss prevention. When combined with the AI platform for retail, companies have the ability to identify at-register theft with real-time video feedback, which alerts staff to intervene or resolves the problem autonomously.

HPE, NVIDIA, and RadiusAI offer comprehensive technologies for accelerating retail analytics. Using video analytics to enhance the customer experience, we help retailers better understand consumer needs and provide efficient customer service to build customer trust and improve brand loyalty. RadiusAI transforms existing video footage into actionable insights, allowing employees to make swift, data-driven decisions and solve in-store problems immediately. With unmatched compute performance and acceleration from HPE and NVIDIA, with industry-leading computer vision from RadiusAI, retailers can turn blind spots into opportunities and keep people at the center of their businesses.

IronYun adds intelligence to retail operation using AI-enabled video analytics. Retailers can turn any IP camera into an AI powerhouse to improve security and track goods at each stage of the supply chain. Layered on top of an accelerated compute platform, IronYun delivers superior accuracy and performance for a variety of use cases (i.e., temperature detection, safety monitoring, and store analytics).

HPE and NVIDIA are working with Deep North to maximize the compute performance of video analytics applications at the edge, highlighting opportunities to influence sales conversion in real time. Deep North makes it faster to unlock new revenue from existing video assets. Their AI-based analytics solution uses deep-learning algorithms to conduct visual analytics, giving retailers the ability to assess, interpret, and predict consumer behavior in physical locations.

Resources

- hpe.com/us/en/solutions/artificial-intelligence.html
- hpe.com/us/en/hpe-ezmeral-data-fabric.html
- hpe.com/us/en/greenlake.html
- nvidia.com/en-us/industries/retail/
- nvidia.com/en-us/data-center/products/ai-enterprise/

With Sensei, we are working to create autonomous stores without checkouts, cashiers, or physical payment for goods. Instead, consumers enjoy queue-free shopping with goods auto-charged to bank cards. This type of environment blends the engagement and customer service benefits of a physical store with the efficiency of a digital store—but it requires processing vast amounts of data at high frame rates. Sensei operates on a cluster of HPE servers and NVIDIA GPUs to run their computer vision tasks and yield point-of-action results.

AiFi is the largest AI platform, empowering retailers to scale autonomous shopping solutions with 100% computer vision. HPE and NVIDIA solutions combined with AiFi technology can support any store format, in a variety of locations. Retailers use AI-enabled computer vision to deliver convenient, accessible, and personalized checkout experiences using cameras only. Consumers enter stores quickly with a credit card or app, and the software recognizes products taken, so consumers can purchase items in-store without having to wait in line to pay.

With Dataiku's visual and collaborative AI platform, retailers have access to easy-to-use analytics tools. Available through the HPE GreenLake Marketplace, Dataiku brings together different types of users, offering both self-service analytics and the operationalization of ML models in production. The joint solution of HPE Ezmeral ML Ops and Dataiku works to systemize AI by enabling everyone on the data team to harness insights using their preferred analytic tools, engines, and interfaces in any operating environment. This way, retailers can focus on high-impact AI projects that generate greater impact.

Pioneering the future of retail

Transformation is critical to navigate the evolving challenges of retail and keep a finger on the pulse of emerging trends. With success at stake, the only solution is to operate faster and smarter. Retailers that invest in intelligence will be able to outpace competitors, predict trends, identify issues that impact profit margins, and rapidly respond to demand.

HPE and NVIDIA are reinventing retail with an AI platform that is designed with today's companies and consumers in mind. We know AI is the key to unlocking new insights and opportunities to deliver superior shopping experiences, optimize business performance, and boost profits. That's why we offer a comprehensive platform that accelerates insight, action, and impact.

We are making innovation faster and simpler. Let HPE and NVIDIA help you prepare for the future of retail. Contact us today.

Learn more at

hpe.com/us/en/solutions/artificial-intelligence/nvidia-collaboration.html

Visit **HPE.com**

Chat now (sales)

© Copyright 2023 Hewlett Packard Enterprise Development LP. The information contained herein is subject to change without notice. The only warranties for Hewlett Packard Enterprise products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. Hewlett Packard Enterprise shall not be liable for technical or editorial errors or omissions contained herein.

NVIDIA and the NVIDIA logo are trademarks and/or registered trademarks of NVIDIA Corporation in the U.S. and other countries. All third-party marks are property of their respective owners.

a00131687ENW

Hewlett Packard Enterprise