Accelerate your AI journey with Red Hat AI

At Red Hat, we offer solutions that help customers innovate so they can achieve their business goals. Red Hat® AI provides a portfolio of solutions that reduce costs and accelerate the development and deployment of artificial intelligence (AI) solutions across hybrid cloud environments. The portfolio includes Red Hat OpenShift® AI, a flexible, scalable AI and machine learning (ML) platform that helps enterprises to create and deliver AI-enabled applications at scale.

Here are 4 ways customers are embracing Red Hat AI:

1 Implementing scalability and flexibility to support generative AI

Delivering an excellent customer experience means providing responsive, highly personalized digital interactions. AI helps organizations to achieve this at scale. However, generative AI (gen AI) is a rapidly evolving technology with substantial data requirements. To keep up, organizations need flexible and scalable options to serve, run, monitor, and fine-tune their gen AI applications on premise, in the cloud, or across a hybrid environment.

Red Hat customers are using OpenShift AI to create knowledge-based applications and chatbot experiences for a wide range of use cases, including online AI-powered text-based sales assistants, customer service chatbots, and code generation to not only automate and supplement code development but also assist developers in code quality assurance.

How Red Hat helps: OpenShift AI offers a single combined AI platform for gen AI, predictive AI, and more traditional application development. It allows organizations to develop, deploy, and manage gen AI applications with greater flexibility and scalability.

With data augmentation through synthetic data generation (SDG), organizations can create additional high quality data that can be used to provide a more accurate model. OpenShift AI also integrates with other solutions, providing organizations with access to large language models (LLMs) and accelerating the delivery of gen AI applications.

2 Providing on-demand infrastructure with ease

To get the best results, data scientists and AI engineers need the freedom to experiment within the constraints of security and compliance requirements. They also need to be able to operate independently, rather than being reliant on other teams to provide the infrastructure and resources they require.

How Red Hat helps: Red Hat customers use OpenShift AI to create self-service environments for data scientists and AI researchers, so they can build experimental models with Jupyter notebooks, TensorFlow, PyTorch, and NVIDIA Graphic Processing Unit (GPU) support without worrying about the underlying infrastructure. The platform simplifies the building, testing, serving, and monitoring of AI models, allowing organizations to standardize AI technologies and processes across teams.

OpenShift AI simplifies the reuse and adaptation of models, enhancing productivity and agility. By empowering employees with controlled autonomy, organizations can accelerate innovation and bring new ideas to market swiftly while maintaining security and compliance.

3 Increasing efficiency in developing predictive AI solutions

AI technology has the power to handle complex analytics problems. It can deliver an automated outcome without the need for time-consuming manual intervention while reducing the risk of human error. To support data-driven or automated decision-making, organizations need to be able to efficiently create predictive AI solutions.

For example, Red Hat customers are using OpenShift AI to rapidly develop AI-powered solutions that:

▸ Automatically classify and route service desk queries to the right team

▸ Quantify financial risk with complex analytical models.

▸ Qualify and prioritize sales leads

▸ Identify healthcare patients at risk of certain conditions

▸ Generate next-best action suggestions for sales teams

How Red Hat helps: With OpenShift AI, organizations can develop predictive AI models with greater efficiency. For example, the solution offers common, open source tooling, allowing teams to standardize while providing the flexibility for users to add their own custom model images and model serving runtimes to meet organizational needs.

By using OpenShift AI to develop predictive AI applications, organizations can also boost their competitive advantage and deliver significant business benefits, including greater employee productivity, customer service quality, and reduced cost and risk.

Learn more about Red Hat OpenShift AI

4 Establishing an MLOps approach

To develop AI applications effectively, organizations need well-defined processes and practices. Many are adopting a machine learning operations (MLOps) approach—a set of workflow practices inspired by DevOps designed to streamline the development and deployment of ML models. MLOps processes help teams work collaboratively to ensure that ML models remain accurate and up to date and model development environments are standardized and automated. To successfully implement MLOps, organizations need the right platform.

How Red Hat helps: OpenShift AI is used by organizations to integrate AI into their existing DevOps workflows. It is optimized for data scientists and developers of AI-enabled applications and provides a fully supported environment to establish MLOps best practices. Data scientists and developers can then rapidly train, deploy, and monitor ML workloads and models on premise and in public clouds.

The AI models built or tuned in OpenShift AI are portable, allowing teams to deploy them on premise and in public cloud or edge environments. It offers a unified platform for managing the entire lifecycle of code, applications, and AI models, from planning and training to tuning and deploying.

Discover how Red Hat can help you benefit from AI

About Red Hat

Red Hat helps customers standardize across environments, develop cloud-native applications, and integrate, automate, secure, and manage complex environments with award-winning support, training, and consulting services.