4 key considerations for implementing AI technology

Business leaders who want to get started with or expand their use of artificial intelligence (AI) technology have common questions and concerns. These include how to use AI to improve operational efficiency and grow their business while maintaining stability and mitigating risk. Here are 4 things to consider, no matter where your company is on its AI implementation journey.

1 Choose a flexible platform

Whether you are just getting started with AI or you want to expand existing use cases and implement AI more broadly, choosing a scalable, flexible platform to support the maturation of your AI deployments is key.

If you are in the initial stages of AI adoption, you need a stable yet flexible and open infrastructure that allows your team to build and run small workloads. The platform you use should support model training at scale as your company's use of AI increases and workloads grow.

Platform stability will allow your data scientists and developers to build models that deliver real value to your customers, who can feel confident that the platform they are using is reliable and focused on security.

If your business is more advanced in its AI capabilities, you need supportive technology that allows your organization to achieve the full potential of AI, which might require a significant investment in graphics processing units (GPUs) and other hardware accelerators.

A scalable, open platform can:

- Provide your company with a good foundation for building and modernizing AI-powered applications in less time.
- Allow you to accelerate the deployment of AI-powered applications.
- Help you gain a competitive advantage in a rapidly changing AI landscape.

With the right infrastructure, your organization will be ready to embrace whatever new AI innovations the future holds.

2 Prepare for hybrid AI

Generative artificial intelligence (gen AI) has increased demand for compute resources and pushed processing to the edge of the network. Gen AI use cases benefit from a mix of cloud and on-premise resources, but some workloads—especially those dictated by compliance issues or data gravity—should only run on premise. With hybrid AI, your team can run workloads wherever it makes the most sense, including closer to where the data is being generated.

Hybrid AI provides several benefits, including:

- **Flexibility.** Running experimental and production workloads in the most suitable environment can create cost savings.
- **Accelerated delivery.** Inference of small workloads at the edge accelerates the delivery of information to users.
- **Reduced risk.** Running AI applications at the edge reduces the need to transfer large amounts of information to the core, resulting in better data protection.

Using hybrid AI, your IT team can train a model in a core datacenter or the public cloud and distribute it to endpoints, where it can infer and provide useful intelligence more quickly. Intensive workloads can be spread across the central cloud and the edge, balancing the need for processing power with the demand for quick and accurate recommendations.

3 Think beyond the first model

Your IT teams can likely put a single AI model into production with minimal effort, but success often depends on the ability to scale. To maximize business value, your teams must be able to deploy multiple versions of that model, perhaps hundreds of times per day.

Implementing machine learning operations (MLOps) processes can help your teams scale deployments with ease. Similar to DevOps, MLOps brings together developers, operations managers, and data scientists to accelerate the development and successful deployment of AI-based applications. With MLOps, your organization can scale application development more easily, allowing you to get more services to your customers in less time.

MLOps can be a cost-effective and efficient means of development. By working together on a common platform, teams can improve efficiencies, potentially leading to time and cost savings that can have a direct effect on your organization's bottom line.

4 Prioritize security

In the face of costly cyber attacks, business executives are increasingly looking for ways to protect sensitive data and mitigate risk. Gen AI creates added urgency, as its use can increase the risk of confidential data exposure.

Maintaining a strong security posture is critical to avoiding threats such as:

- **Data poisoning.** This injects incorrect or malicious data into AI models, resulting in inaccurate information or fraudulent recommendations.
- **Model theft.** This is the reverse engineering of a machine learning model and extraction of data using the copy.
- **Backdoor attacks.** This is the hiding of malicious code in an AI model, thereby allowing the attacker to steal information.

These and other tactics can pose a significant threat to your business, exposing your organization to both financial and legal risks. Minimize the vulnerabilities associated with AI with help from a trusted application platform provider that has a proven track record and established processes for handling security patches and releases.

Learn more

Accelerate your AI strategy. Explore Red Hat's AI solutions.

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.

North America	Europe, Middle East, and Africa	Asia Pacific	Latin America
1 888 REDHAT1 www.redhat.com	00800 7334 2835 europe@redhat.com	+65 6490 4200 apac@redhat.com	+54 11 4329 7300 info-latam@redhat.com

facebook.com/redhatinc
twitter.com/RedHat
linkedin.com/company/red-hat

redhat.com