

# Quatro considerações importantes sobre a implementação da tecnologia de IA

Os líderes empresariais que querem começar a usar a inteligência artificial (IA) ou expandir o uso dessa tecnologia têm dúvidas e preocupações em comum. Isso inclui como utilizar a IA para melhorar a eficiência operacional e impulsionar os negócios, mantendo a estabilidade e reduzindo os riscos. Confira a seguir quatro considerações a serem feitas, seja qual for o estágio da sua empresa na jornada de implementação da IA.

## 1 Escolha uma plataforma flexível

Mesmo que você esteja apenas começando a utilizar a IA ou queira expandir os casos de uso e implementar a tecnologia de forma mais ampla, é essencial escolher uma plataforma escalável e flexível para viabilizar o amadurecimento das implantações de inteligência artificial.

Se você está nos primeiros estágios da adoção da IA, é necessário adotar uma infraestrutura aberta e flexível para que a equipe possa criar e executar cargas de trabalho pequenas. A plataforma implementada precisa oferecer suporte ao treinamento de modelos em escala à medida que o uso da IA aumenta na sua empresa e as cargas de trabalho crescem.

Com estabilidade na plataforma, os cientistas de dados e os desenvolvedores podem criar modelos que agregam valor real aos clientes. Além disso, eles vão ter a certeza de que a plataforma que usam é confiável e voltada à segurança.

Se sua empresa é mais madura em relação aos recursos de inteligência artificial, você precisa de uma tecnologia complementar para que a organização atinja todo o potencial da IA. Isso pode exigir um investimento significativo em unidades de processamento gráficos (GPUs) e em outros aceleradores de hardware.

Com uma plataforma aberta e escalável, você pode:

- ▶ Oferecer à sua empresa uma ótima base para criar e modernizar aplicações com tecnologia de IA em menos tempo.
- ▶ Acelerar a implantação de aplicações com tecnologia de IA.
- ▶ Assegurar uma vantagem competitiva em um cenário de IA que muda rapidamente.

Com a infraestrutura certa, sua organização se prepara para adotar as inovações de IA que surgirem no futuro.

## 2 Faça sua preparação para a IA híbrida

A inteligência artificial generativa (gen AI) aumentou a demanda por recursos de computação e levou o processamento para a edge da rede. A combinação de recursos on-premises e de nuvem pode ser boa para os casos de uso da gen AI. No entanto, é necessário executar apenas on-premises algumas cargas de trabalho, principalmente aquelas com questões de conformidade ou gravidade dos dados. Com a IA híbrida, sua equipe pode executar cargas de trabalho onde mais apropriado, incluindo mais perto de onde os dados estão sendo gerados.

A IA híbrida oferece vários benefícios, incluindo o seguinte:

- ▶ **Flexibilidade:** executar cargas de trabalho experimentais ou de produção no ambiente mais adequado produz economias de custo.
- ▶ **Entrega mais rápida:** a inferência de cargas de trabalho pequenas na edge acelera a entrega de informações aos usuários.
- ▶ **Menos riscos:** a execução de aplicações de IA na edge diminui a necessidade de transferir grandes volumes de informações para o núcleo, o que aprimora a proteção dos dados.

Com o uso da IA híbrida, sua equipe de TI pode treinar um modelo no data center principal ou na nuvem pública, além de realizar a distribuição dele para os endpoints. Lá, o modelo pode fazer inferências e oferecer informações úteis com mais rapidez. É possível espalhar as cargas de trabalho exigentes pela nuvem central e pela edge. Assim, você encontra um equilíbrio entre a capacidade de processamento necessária e a demanda por recomendações rápidas e precisas.

### 3 Pense além do primeiro modelo

É provável que suas equipes de TI consigam colocar um modelo de IA em produção com pouco esforço. No entanto, o sucesso muitas vezes depende da escalabilidade. Para maximizar o valor de negócio, suas equipes precisam ser capazes de implantar várias versões desse modelo, talvez até mesmo centenas de vezes por dia.

Com a implementação de processos de operações de machine learning ([MLOps](#)), suas equipes podem escalar as implantações com facilidade. Assim como o DevOps, o MLOps une desenvolvedores, gerentes de operações e cientistas de dados para acelerar a criação e a implantação de aplicações baseadas em IA. Com o MLOps, sua organização pode escalar o desenvolvimento de aplicações com mais facilidade, o que acelera a entrega de mais serviços para os clientes.

O MLOps é uma forma econômica e eficiente de desenvolvimento. Ao trabalharem juntas em uma mesma plataforma, as equipes aumentam a eficiência e podem produzir economias de tempo e custo que têm um impacto direto nos resultados financeiros da organização.

### 4 Priorize a segurança

Diante de ciberataques caros, os executivos empresariais estão cada vez mais em busca de maneiras de proteger os dados confidenciais e reduzir os riscos. Isso se torna ainda mais urgente devido à gen AI, já que o uso dessa tecnologia pode aumentar o risco de exposição de dados confidenciais.

Manter uma forte postura de segurança é essencial para evitar ameaças como estas:

- ▶ **Envenenamento de dados:** são injetados dados incorretos ou maliciosos nos modelos de IA, o que produz informações imprecisas ou recomendações fraudulentas.
- ▶ **Roubo de modelo:** isso se refere à engenharia reversa de um modelo de machine learning e à extração dos dados usando a cópia dele.
- ▶ **Ataques de backdoor:** código malicioso é escondido no modelo de IA, permitindo assim que o invasor roube informações.

Esses e outros métodos de ataque são ameaças significativas para sua empresa que expõem a organização a riscos financeiros e jurídicos. Diminua as vulnerabilidades associadas à IA com a ajuda de uma provedora confiável de plataformas de aplicações que tem um histórico de sucesso comprovado e processos estabelecidos para lidar com patches de segurança e lançamentos.

---

#### Mais informações

Acelere sua estratégia de IA. [Explore](#) as soluções de IA da Red Hat.



#### Sobre a Red Hat

A Red Hat ajuda os clientes a definir padrões entre diferentes ambientes e a desenvolver aplicações nativas em nuvem, além de integrar, automatizar, proteger e gerenciar ambientes complexos com serviços de consultoria, treinamento e suporte [premiados](#).

---

#### AMÉRICA LATINA

+54 11 4329 7300  
latammktg@redhat.com

#### BRASIL

+55 11 3629 6000  
marketing-br@redhat.com

[f](#) facebook.com/redhatinc  
[X](#) @redhatbr  
[in](#) linkedin.com/company/red-hat-brasil