

# Como implementar modelos personalizados de IA generativa de uma forma mais inteligente



## Está na hora de ser realista

A Inteligência Artificial Generativa (GenAI) pode gerar vantagens transformadoras em todos os setores, mas as organizações precisam se preparar. O entusiasmo e os investimentos potencialmente sem direção devem ser canalizados em ações e resultados focados.

E aqueles CIOs que sabem da necessidade de priorizar a GenAI agora estão considerando serviços de consultoria, profissionais e gerenciados como um próximo passo essencial para alcançar resultados reais.

O ponto de partida deve ser obter uma compreensão profunda da sua prontidão em IA. Sua infraestrutura tem o tamanho necessário para executar modelos de GenAI? Ela é flexível o suficiente para atender requisitos em constante mudança? Seu conjunto de dados está bem estruturado para os casos de uso prioritários? Você considerou os principais desafios que pode enfrentar em encontrar as habilidades certas e orçamentos?

**43%**

As tecnologias emergentes estão no topo das prioridades dos CIOs, com 43% sentindo "pressão urgente" para tratar a inteligência artificial/aprendizado de máquina como prioridade<sup>1</sup>.

**80%**

Sentem-se significativamente ou moderadamente desafiados em desenvolver soluções de IA que entrem em operação rapidamente<sup>2</sup>.

**83%**

A porcentagem de CIOs preocupados com a falta de recursos suficientes para investir efetivamente em inovação para a infraestrutura de TI da organização<sup>3</sup>.

### Parceria poderosa: por que Lenovo e NVIDIA?

Acreditamos que alianças são a maneira mais poderosa de impulsionar a inovação. A colaboração permite abordar com mais eficiência os principais desafios do setor e desbloquear novas oportunidades. É por isso que Lenovo e NVIDIA forjaram nossa parceria estratégica. Entendemos que todas as necessidades são diferentes, e nossa parceria ajuda organizações em qualquer etapa da transformação.

Os serviços de IA da Lenovo fornecem uma solução de fornecedor único para guiar a implementação e implantar facilmente soluções personalizadas. Aproveitando a NVIDIA Enterprise AI e o nosso ecossistema de ponta de Inovadores de IA, juntos oferecemos às empresas uma plataforma que impulsiona a diferenciação, promove o crescimento e ajuda a moldar o futuro das soluções impulsionadas por IA.



<sup>1</sup> Lenovo, Global Study of CIOs, 2023.

<sup>2</sup> Lenovo, State of Enterprise AI Buying, Quantitative Survey, 2024.

<sup>3</sup> Lenovo, Global Study of CIOs, 2023.

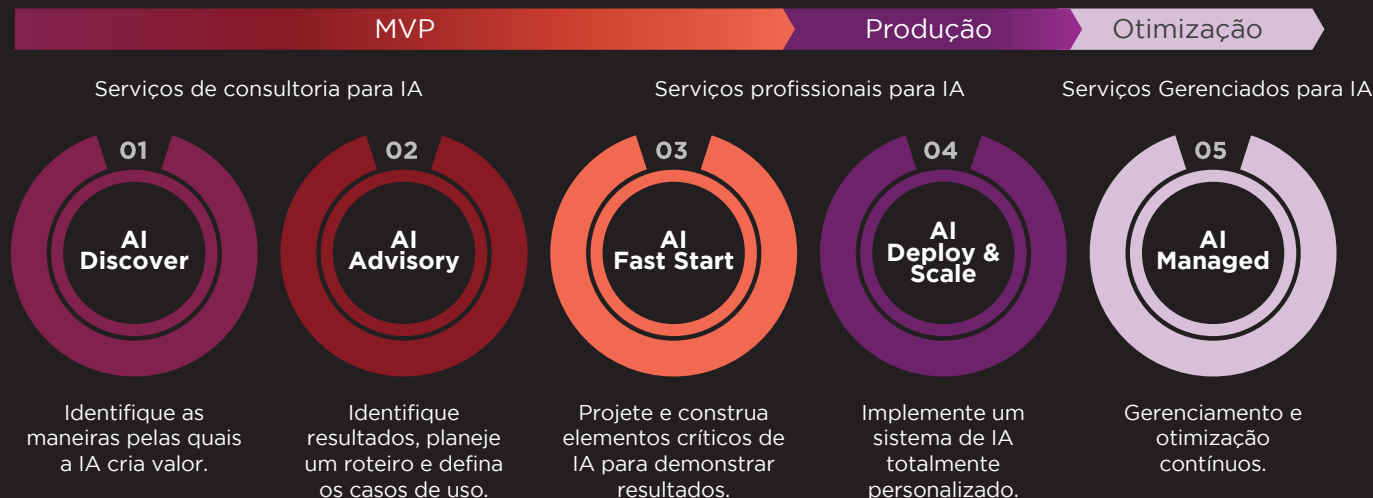
## Os desafios que você enfrenta e os benefícios que entregamos

Os principais desafios que você enfrenta	Os principais benefícios da nossa solução conjunta
<b>Quantidade e qualidade de dados</b> Para que os modelos de GenAI compreendam, prevejam e gerem texto humano, você precisa de dados diversos, de alta qualidade e de alto volume – mas as organizações podem ter dificuldades para coletar e organizar seus dados.	Avaliação de como seus dados são mais bem utilizados para os resultados que você busca é uma parte essencial da nossa avaliação única de prontidão em IA. Nossa solução com NVIDIA foca na personalização de modelos pré-treinados, o que requer significativamente menos dados do que treinar do zero.
<b>Confiança</b> Trabalhar com terceiros para IA inevitavelmente significa entregar dados – seus ou de seus clientes. A segurança dos dados é claramente uma preocupação.	A Lenovo é segura por design, com certificação ISO 27001 e um programa de fornecedores confiáveis para garantir que todos os nossos parceiros atendam nossos altos padrões de segurança. Já gerenciamos dados e dispositivos para muitos de nossos clientes no mundo todo, mantendo conformidade com políticas e leis em constante mudança.
<b>Aceleração de tempo para valor e escalabilidade</b> Muitas organizações descobrem que os pilotos de GenAI ficam presos em estágios de prova de conceito e carecem de escalabilidade ou versatilidade para serem implementados em todas as funções.	Nossa solução aproveita o NVIDIA NIM, um conjunto de microserviços fáceis de usar, projetados especificamente para acelerar o tempo de solução e a implementação de modelos fundamentais. Também oferecemos softwares verticais específicos de GenAI de ISVs em nosso programa AI Innovators.
<b>Infraestrutura de GenAI</b> A escala de cálculos necessária para treinar modelos GenAI do zero é imensa. Você precisa de uma infraestrutura robusta e de grande escala, que é complexa e cara de gerenciar.	Os requisitos computacionais para personalizar um modelo pré-treinado são significativamente menores. Nosso foco está na personalização de Modelos de Linguagem (LMs) supervisionados e pré-treinados.
<b>Treinamento e software</b> Implementar GenAI em aplicações exige o hardware certo, juntamente com uma pilha de software e ferramentas que superem os obstáculos de treinamento e inferência – de desenvolvimento de algoritmos a aceleração de inferência em infraestrutura distribuída.	Nossa solução utiliza toda a Plataforma de IA da NVIDIA e nosso IP próprio Lenovo para superar os obstáculos de treinamento e inferência. Com as GPUs NVIDIA mais recentes e sistemas Lenovo Certificados NVIDIA, o treinamento e a implementação são acelerados.
<b>Riscos</b> As implementações de GenAI podem ser inerentemente arriscadas – superestimar os recursos necessários compromete o ROI; subestimar pode resultar em falha do modelo.	Lenovo TruScale e nossos serviços gerenciados de IA oferecem flexibilidade de gerenciamento contínuo como um serviço e otimização do seu sistema personalizado de GenAI.
<b>Habilidades</b> Para construir e ajustar modelos generativos, você precisa adicionar novas habilidades técnicas à sua equipe de ciência de dados, IA ou desenvolvimento de aplicações – ou terceirizar tarefas para que suas equipes possam focar em outras áreas críticas.	Os melhores cientistas de dados da Lenovo, a propriedade intelectual pioneira da NVIDIA e nosso ecossistema de Inovadores de IA de ponta ajudam a levar sua equipe a resultados elevados com GenAI.

## Como levamos você a um sistema de IA personalizado, pronto para produção e totalmente gerenciado

### Um portfólio abrangente de serviços de IA

Caminho para um sistema de IA personalizado, totalmente gerenciado e pronto para produção



Sua jornada para concretizar sua visão de IA começa com o nosso workshop AI Discover, conduzido pelo Centro de Excelência em IA da Lenovo. O AI Discover identifica os desafios específicos que você enfrenta e define como a IA pode ser aproveitada não apenas para superá-los, mas também para garantir que você atinja novos e melhores resultados.

O que você deve priorizar? O que se encaixa nas necessidades do seu negócio? O que é seguro e escalável? Ajudamos você a criar uma visão clara para a IA em seu negócio - do planejamento à implementação.

O workshop pode direcionar você para nosso GenAI Advisory com NVIDIA, AI PC Advisory, Digital Workplace AI Advisory, Business AI Advisory ou ESG Advisory. Nos encontramos onde você está, oferecendo personalização, incluindo a logística e o local onde conduzimos o workshop, a duração do processo e os custos envolvidos.



#### A partir desse estágio inicial do AI Discover, avançamos passo a passo:

##### AI Advisory

Define seu roteiro e articula seu caso de negócios para GenAI. Você precisa de uma estratégia bem estruturada e um plano para concretizar sua visão de IA, focando em casos de uso de alta prioridade. Criamos um plano de curto e longo prazo com próximos passos detalhados.

##### AI Fast Start

O AI Fast Start constrói uma solução personalizada baseada nos casos de uso necessários. Você terá uma prova de conceito poderosa que roda nos seus dados, com nossa equipe de especialistas demonstrando o valor da IA para o seu negócio.

Escolha entre três caminhos para sua solução personalizada de IA:

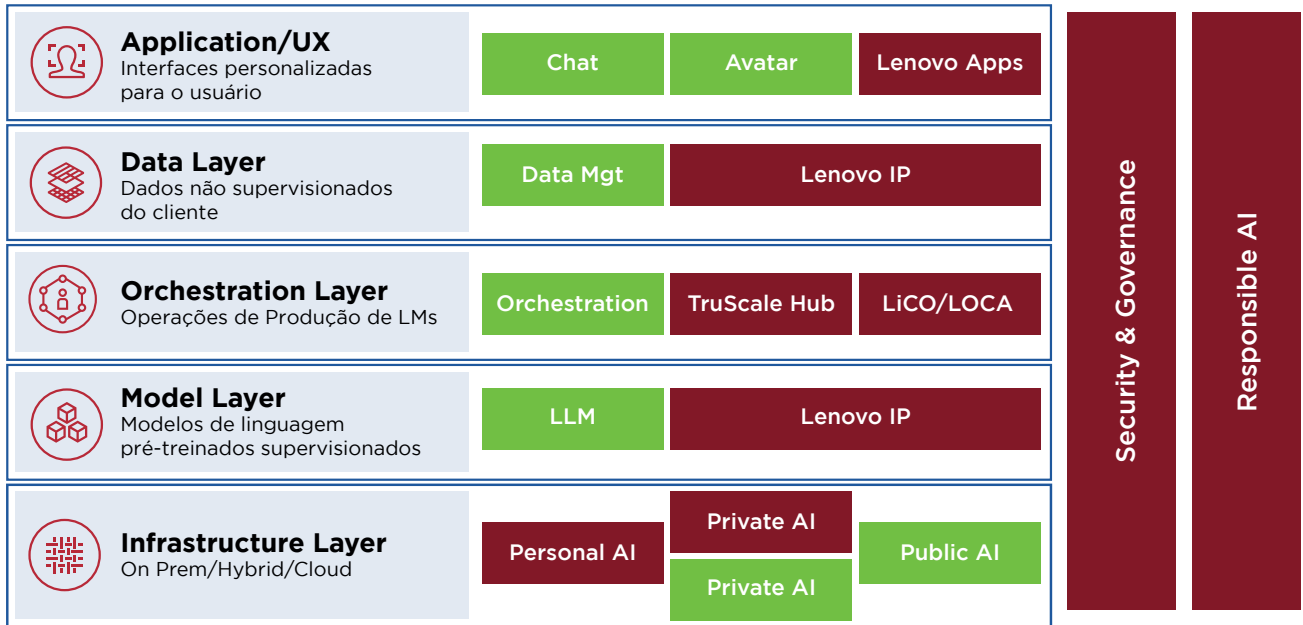
- AI Enterprise: construa soluções personalizadas de IA usando seus dados em GPUs NVIDIA e o software NVIDIA AI Enterprise.
- AI Innovators: entregue uma solução validada de terceiros otimizada para IA Lenovo e hardware de edge.
- NVIDIA NIM: implemente micros serviços prontos para uso com o NVIDIA NIM, os micros serviços containerizados da NVIDIA.

##### AI Deploy & Scale

Escalamos a solução para toda a sua empresa, implementando-a em seu data center ou na nuvem. Continuamos otimizando os sistemas e desenvolvendo habilidades para garantir o sucesso contínuo. Nossos consultores, assessores e cientistas de dados são dedicados a impulsionar você com orientações especializadas sobre como manter sua infraestrutura atualizada e de ponta. Assim, você está sempre operando com desempenho máximo.

##### AI Managed

Implementamos sua abordagem gerenciada de ponta a ponta para sistemas de IA, com custos disponíveis como assinatura através do Lenovo TruScale. À medida que a IA evolui, suas prioridades evoluem junto. Ajudamos a planejar além das iniciativas iniciais e a construir uma cultura de inovação, usando a IA como catalisador para melhorias contínuas.



**Aqui estão os principais pontos sobre como revelamos o potencial da nossa GenAI:**

- Versão privada e dedicada do ChatGPT que você pode ter com segurança - ela roda no seu data center sob seu controle, supervisionada pela Lenovo.
- Modelo personalizado para cada um dos nossos clientes, adaptado às suas necessidades únicas.
- Large Language Model (LLM) com múltiplas interfaces de linguagem para chat e tradução ao vivo.
- Utiliza a melhor pilha de software da NVIDIA, o framework NVIDIA NeMo™ GenAI, e o IP da Lenovo.

O NVIDIA NeMo é um framework em nuvem nativo de ponta a ponta para construir, personalizar e implementar modelos de GenAI, em qualquer lugar. Ele fornece uma solução completa em todo o pipeline LLM, desde o processamento de dados até a inferência. O NeMo permite que você treine, personalize e implemente LLMs em escala, reduzindo o tempo para solução e aumentando o seu ROI.

Com o suporte da nossa prática de Serviços de IA, nossa GenAI permite que você utilize uma abordagem híbrida de nuvem. Você pode criar modelos de IA personalizados usando o serviço em nuvem NVIDIA AI Foundations e, em seguida, executá-los em qualquer lugar que desejar, com sistemas Lenovo locais equipados com o mais recente hardware e software projetados para GenAI.

**Responsável e seguro**

Nosso foco é impedir alucinações e garantir o uso ético e responsável de IA, apoiados pelo Comitê de Responsabilidade de IA da Lenovo. Estamos comprometidos em garantir IA segura - nossa solução não expõe os dados ao público como fariam outros LLMs de código aberto.

**Nossas principais capacidades para o sucesso de sua empresa**

Independentemente de onde você estiver na sua jornada com GenAI, a Lenovo e a NVIDIA podem ajudar você a alcançar melhores resultados mais rapidamente com nossa abordagem única. Juntos, podemos evitar atrasos e ir direto aos resultados com nossos serviços abrangentes, personalizados e confiáveis.





### Indo além da tecnologia

Avaliamos pessoas, processos, tecnologia e segurança para obter uma visão completa do seu negócio.

### Jornadas simplificadas para clientes

Aproveitamos nossas parcerias para reduzir riscos e simplificar sua jornada.



### Especialização de GenAI de ponta

Você trabalhará perfeitamente com consultores experientes da Lenovo e engenheiros, além de pioneiros de GenAI na NVIDIA. Executamos GenAI em nossos próprios sistemas há anos, em toda nossa cadeia de suprimentos e nosso software de e-commerce, no qual vendemos cinco dispositivos a cada segundo. Aproveitamos o poder da GenAI para permitir que nossos desenvolvedores gerem novo código mais rápido, entregando melhores resultados para nossos clientes – sempre garantindo sucesso seguro e responsável.



### Pronto para produção

Criamos modelos personalizados de nível corporativo e implementamos infraestrutura que executa cargas de trabalho de GenAI com segurança e de forma otimizada.



### Segurança confiável

Aplicamos os mais altos padrões e os controles de segurança mais rígidos para garantir que você use GenAI com segurança.



### Ecosistema amplo

Oferecemos flexibilidade por meio de uma ampla e experiente rede de parceiros AI Innovator, garantindo facilidade de implementação com soluções desenvolvidas para o hardware Lenovo.



### Fluxos de trabalho acelerados

Proporcionamos desempenho poderoso na borda com sistemas Lenovo Certificados NVIDIA para soluções edge.



### Implementação rápida

Fazemos tudo rapidamente, ajudando você a ir da estratégia à implementação em 2-3 meses.



### GenAI responsável

Nosso Comitê de Responsabilidade de IA aborda uma ampla gama de desafios no espaço de GenAI, com base em 6 pilares: Diversidade & Inclusão, Privacidade & Segurança, Responsabilidade & Confiabilidade, Explicabilidade, Transparência e Impacto Ambiental & Social.

Pronto para uma implementação mais inteligente de modelos personalizados de GenAI?

Entre em contato para agendar hoje mesmo o seu workshop AI Discover.



A visão é sua.  
Alcance com Lenovo  
e NVIDIA.

Smarter  
technology  
for all

Lenovo