DELL Technologies

NVIDIA

As cinco principais considerações de segurança para IA generativa (GenAI)

Acelere a adoção de uma base de infraestrutura segura e escalável com o Dell AI Factory with NVIDIA

O potencial transformador da IA generativa

A IA generativa tem potencial para mudar o jogo de maneiras que os visionários estão apenas começando a imaginar.

76%

dos líderes empresariais e de TI acreditam que a IA generativa proporcionará valor transformador à sua organização.[1]

IA

Análises avançadas e técnicas baseadas em lógica para interpretar eventos, dar suporte e automatizar decisões e ações.

IA generativa

Tecnologias e técnicas que aproveitam grandes volumes de dados para gerar novo conteúdo a partir de prompts de linguagem natural ou outras entradas não codificadas e não tradicionais.

Simulação

- Gêmeo digital
- Dados sintéticos
- Estruturas de design
- Previsões

Criação de conteúdo

- Criação de códigos
- Matemática
- Escrita/fala
- Imagem/vídeo
- Áudio

Descoberta de conteúdo

- Pesquisa de linguagem natural
- Análises de conjunto de dados de grande porte
- Gerenciamento de conhecimento
- Treinamento e educação personalizados

Experiência do usuário

- Traduções em tempo real para mais de 70 idiomas
- Interações personalizadas usando expressões faciais naturais e linguagem corporal

1 Estudo Innovation Catalyst da Dell Technologies, fevereiro de 2024

Maior potencial, maior risco

Os líderes de negócios tendem a querer agir rapidamente, ignorando as implicações que envolvem dados, conformidade, governança e outros riscos. Mas a IA generativa é uma faca de dois gumes quando se trata de segurança.

Benefícios

- Detecção de ameaças aprimorada
- Melhor eficiência operacional
- Treinamento em conscientização de segurança personalizada

Desvantagens

- Maior sofisticação de ataques
- Engenharia social avançada
- Shadow AI

33%

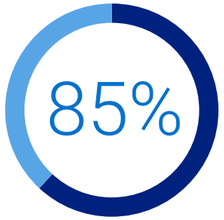
dos entrevistados listaram a segurança cibernética como o principal risco de IA generativa que suas organizações estão trabalhando para reduzir.[2]

2 Pesquisa global da McKinsey sobre IA: O estado da IA no início de maio de 2024

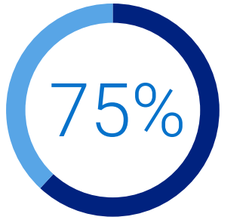
CONSIDERAÇÃO 1

O novo cenário de ameaças

Junto com a promessa da IA generativa, existe uma realidade preocupante: os invasores estão criando ataques novos e mais intrincados que podem contornar as defesas convencionais, tornando difícil para as equipes de segurança cibernética acompanharem o ritmo.



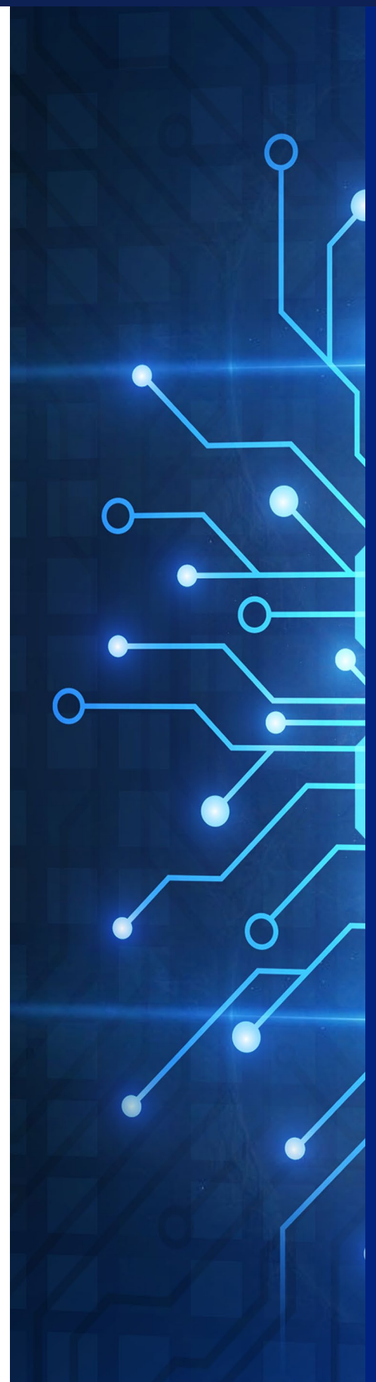
dos entrevistados acreditam que a IA tornou os ataques à segurança cibernética mais sofisticados.[3]



dos profissionais de segurança observaram um aumento nos ataques nos últimos 12 meses.[4]

Para se proteger contra essas ameaças emergentes, as empresas devem se concentrar em minimizar a superfície de ataque por meio de testes de penetração, monitoramento e auditoria, por exemplo.

3 Pesquisa sobre riscos humanos em segurança cibernética de 2024, EY, maio de 2024
4 Relatório Voice of SecOps: "IA generativa e segurança cibernética: Futuro brilhante ou campo de batalha de negócios?" 2023



Vetores de ataque emergentes

Malware avançado
Malware cada vez mais sofisticado que usa IA generativa para "evoluir automaticamente", mudando continuamente seu código para passar despercebido pela segurança existente, como a detecção baseada em assinatura.

E-mails e campanhas de phishing altamente personalizados
Frequência crescente de e-mails maliciosos com aparência autêntica sem sinais de fraude usuais.

Convencimento de dados deep fake
Roubo de identidade, fraude financeira e desinformação facilitados pela capacidade de imitar ações humanas, como escrita, fala, imagens ou vídeo.

Reconhecimento automatizado
Coleta de informações que identificam vulnerabilidades e pontos fracos na rede ou no sistema de um alvo potencial para facilitar ataques mais direcionados.

CONSIDERAÇÃO 2

Riscos de implantação e implementação

As organizações que desejam aproveitar os possíveis benefícios da IA generativa precisam de grandes quantidades de dados de alta qualidade, entradas que os modelos podem usar para produzir os melhores resultados. Mas dados e riscos andam lado a lado. Antes de aproveitar qualquer informação, as empresas devem avaliar cuidadosamente e levar em conta seus requisitos, entradas e riscos exclusivos.

Vulnerabilidades dos Modelos de linguagem grandes (LLMs)

Os serviços da IA generativa são vulneráveis a ataques de injeção imediata, em que os invasores manipulam os resultados para contornar as grades de proteção de segurança ou obter acesso não autorizado a arquivos que podem ter sido usados para refinar o modelo.

Envenenamento de dados

Os invasores podem deliberadamente alimentar dados alterados em um LLM durante a fase de treinamento. Isso pode tornar o modelo vulnerável a ataques por meio de backdoors incorporados aos dados. Um exemplo do mundo real é atacar e explorar filtros de spam, treinando-os em e-mails de spam.

Complexidade regulamentar

Reguladores em todo o mundo estão correndo para entender, controlar e garantir a segurança da IA generativa. Embora os modelos de IA generativa estejam sujeitos às regras atuais de soberania de dados que ditam como os dados são armazenados, processados e usados, os órgãos governamentais ainda estão definindo a supervisão de informações de IP e direitos autorais. A adesão às normas pode custar caro, mas o não cumprimento das normas estabelecidas e emergentes pode resultar em multas e outras penalidades.

CONSIDERAÇÃO 3

Shadow AI

Hoje, muitos funcionários já usam geradores públicos de texto, imagem e vídeo, como o ChatGPT, para aumentar seus fluxos de trabalho diários. No entanto, quando essas ferramentas são usadas sem governança adequada, elas representam uma ameaça crítica para as organizações que tentam proteger a propriedade intelectual e os dados corporativos. Esse uso não autorizado da IA generativa é conhecido como IA invisível.

Perda de propriedade intelectual

As empresas já estão lidando com a perda de propriedade intelectual de funcionários que compartilham informações confidenciais em ferramentas públicas de IA generativa.

Vazamento de dados do código-fonte

Os desenvolvedores que tentam otimizar o código-fonte usando o ChatGPT causam vazamento de dados.

Para enfrentar os desafios da IA invisível, as empresas devem implementar um conselho ou diretoria em toda a empresa com autoridade para tomar decisões que envolvam governança segura de IA.

Onde os dados residem? Onde as cargas de trabalho devem ser colocadas?

A IA funciona melhor quando combinada com seus dados, onde quer que eles estejam. Com controle total sobre a infraestrutura e LLMs, não há risco de perda de IP ou vazamento de dados do código-fonte.

Custos

O aproveitamento das implementações no local pode reduzir o TCO em até 75% em 3 anos.[5º]

Segurança e privacidade

Crie ambientes seguros de IA/IA generativa em toda a organização com fluxos de trabalho e operações no local. Exerça controle rigoroso sobre a segurança de dados e siga as normas de conformidade, especialmente em setores que lidam com dados confidenciais.

5 Com base em um estudo do Enterprise Strategy Group encomendado pela Dell, que compara a infraestrutura local da Dell com a infraestrutura em nuvem pública nativa "as a service", abril de 2024. Os modelos analisados mostram que um LLM com 7 bilhões de parâmetros utilizando o RAG oferece a uma organização de 5.000 usuários uma economia de até 38%, enquanto um LLM com 70 bilhões de parâmetros utilizando o RAG oferece a uma organização de 50.000 usuários uma economia de até 75%. Os resultados reais podem variar. Resumo econômico

CONSIDERAÇÃO 4

Critérios de avaliação

No último ano, a comunidade de IA tem se concentrado cada vez mais em três questões principais: desenvolvimento e implantação responsáveis, avaliação do impacto e redução de riscos. À medida que as empresas avaliam os modelos de IA generativa, elas devem levar em conta algumas restrições importantes:

Ausência de requisitos de geração de relatórios consistentes

Os principais desenvolvedores testam principalmente seus modelos em relação a diferentes referências de desempenho responsáveis de IA. Devido a essa significativa falta de padronização nos relatórios, é difícil comparar metodicamente os riscos e as limitações dos principais modelos de IA.

Material protegido por direitos autorais nos resultados

Os resultados de LLMs populares podem conter material protegido por direitos autorais, violando potencialmente a lei e criando risco de penalidade para as empresas que usam o material.

As vulnerabilidades estão cada vez mais complexas

Os pesquisadores estão encontrando estratégias menos óbvias que fazem com que os LLMs exibam comportamento prejudicial, como pedir aos modelos que repitam infinitamente palavras aleatórias.

Ausência de transparência dos desenvolvedores

Em muitos casos, os desenvolvedores de IA não estão cientes de seus dados e metodologias de treinamento. Isso dificulta os esforços para entender melhor a eficiência e a segurança dos sistemas de IA.

CONSIDERAÇÃO 5

Benefícios de segurança

Paralelamente aos riscos de segurança da IA generativa, estão seus potenciais benefícios de segurança. A IA generativa está se tornando um aliado crucial na segurança cibernética, abrindo novos caminhos de proteção.

Agora você pode começar a criar operações de segurança escaláveis com acesso mais rápido a insights mais avançados e detecção automática de ameaças, proporcionando eficiência e complementando equipes de segurança com equipes insuficientes.

Detecção e resposta a ameaças

Ao analisar dados históricos e identificar padrões e anomalias, a IA generativa pode reconhecer ameaças novas e em evolução em tempo real. Ela pode monitorar continuamente o tráfego de rede, os logs do sistema e o comportamento do usuário e identificar prontamente atividades irregulares que possam significar ameaças à segurança.

O resultado é uma detecção de ameaças poderosamente adaptável, permitindo uma resposta rápida aos vetores de ataque em constante mudança e fornecendo um mecanismo de defesa proativo contra ameaças cibernéticas emergentes.

Simulação e treinamento contra ameaças

Com a IA generativa, as empresas podem simular uma ampla gama de ameaças e cenários de ataque à segurança cibernética em um ambiente controlado. Como resultado, as equipes ficam mais bem preparadas para identificar, reduzir ameaças cibernéticas e responder a elas quando o tempo for essencial.

Análise detalhada e resumo

A IA generativa capacita as equipes a investigar dados de diferentes fontes ou módulos, permitindo que elas realizem análises de dados tradicionalmente demoradas e tediosas com mais rapidez e mais precisão. As equipes também podem criar resumos em linguagem natural de incidentes e avaliações de ameaças, melhorando a eficiência e aumentando o resultado da equipe.

Treinamento em conscientização de segurança personalizada

Ao incorporar a IA de conversação à IA generativa e incorporar um avatar de IA à interface do usuário, as organizações podem oferecer interações personalizadas (disponíveis em escala 24x7) usando expressões faciais naturais e linguagem corporal. Isso pode ser usado para treinamento e educação em segurança, proporcionando uma experiência de aprendizado mais natural, personalizada e interativa, avaliações automatizadas e muito mais.

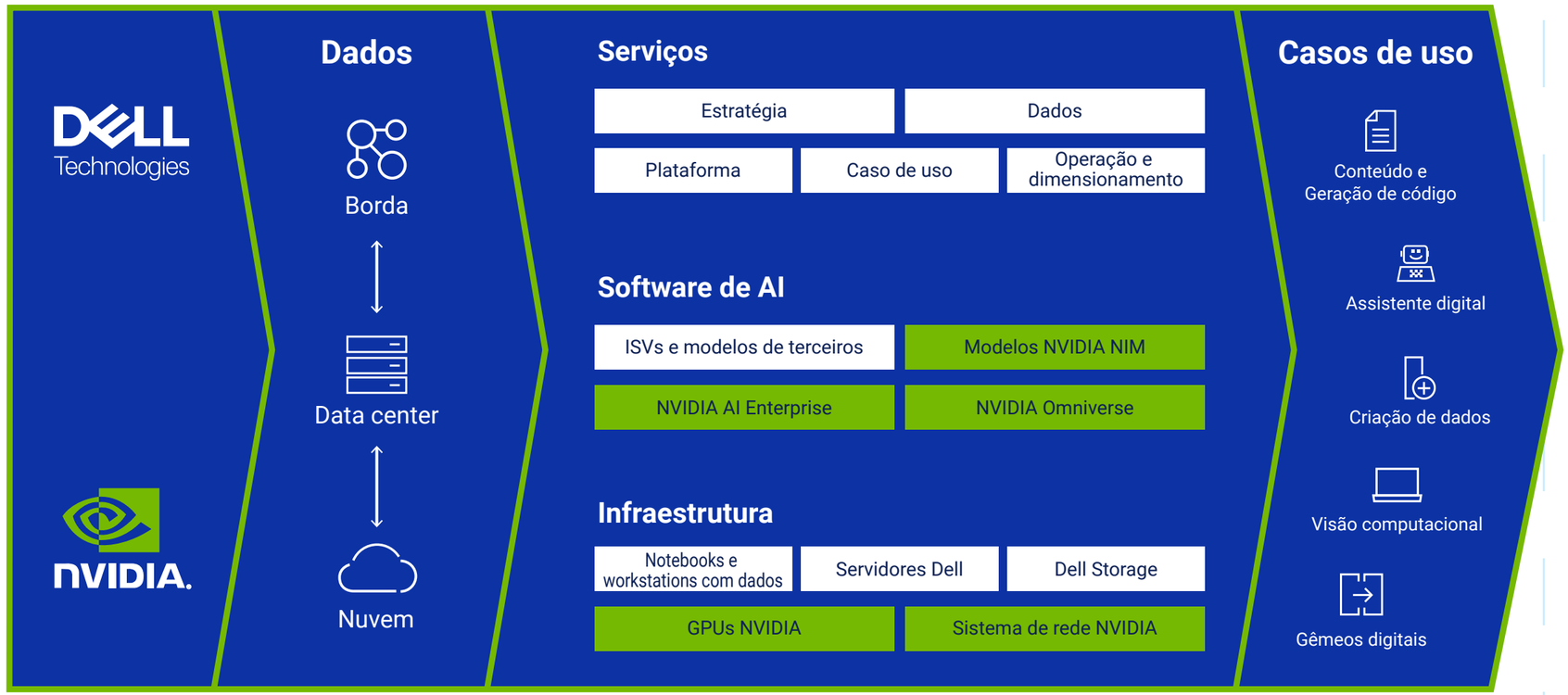
Dell AI Factory with NVIDIA

Acelere sua jornada de IA e transforme com segurança seus dados em insights com a primeira solução de IA abrangente e turnkey do setor. O Dell AI Factory with NVIDIA atende às necessidades complexas das empresas que buscam aproveitar a IA e a IA generativa. Com infraestrutura e serviços líderes do setor, juntamente com o software de IA NVIDIA, você pode aumentar o time-to-value de seus projetos simplificando o desenvolvimento e a implementação.

- Reduza o risco de comprometimento com a infraestrutura que apresenta segurança intrínseca, inclusive raiz de confiança e outros recursos importantes.
- Proteja seus dados contra vazamentos que podem resultar em perda de propriedade intelectual com uma solução de IA no local que você controla.
- Atenda aos rigorosos requisitos de conformidade e soberania de dados trazendo a IA para seus dados com acesso seguro.
- Proteja a privacidade da parte interessada controlando onde e quem tem acesso aos seus dados.

Dell AI Factory with NVIDIA

PRIMEIRA SOLUÇÃO COMPLETA DE IA EMPRESARIAL DO SETOR



Os dados alimentam a fábrica de IA e seus casos de uso

Os dados mais valiosos estão no local e na borda. A Dell Technologies ajuda a incorporar a IA nesses dados valiosos e é líder em armazenamento, proteção e gerenciamento desses dados.

Caso de uso para resultados

A fábrica de IA produz resultados para os negócios baseados nos seus casos de uso de maior prioridade. A Dell Technologies simplifica a implementação dos casos de uso de IA mais importantes com soluções validadas e serviços personalizados.

Não deixe que os riscos de segurança impeçam a inovação

Deixe-nos ajudá-lo a navegar pelo mundo da IA e da IA generativa para que você possa aproveitar as recompensas.

PLANEJAMENTO DE ESTRATÉGIA

Accelerator Workshop for GenAI gratuito

- Inicie sua jornada para desenvolver uma estratégia vencedora
- Abordar desafios e lacunas, priorizar objetivos e identificar oportunidades
- Obtenha uma avaliação de prontidão para se aprofundar nos requisitos de infraestrutura, modelos de IA, integrações operacionais e muito mais

PREPARO TÉCNICO

Laboratório móvel pronto para uso

Dê início à jornada rumo ao sucesso. Inclui uma workstation móvel Dell Precision 5690/7780 com GPUs NVIDIA e dois dias de serviços de consultoria para ajudá-lo a começar.

- Ambiente de área restrita portátil para testes e demonstração de IA generativa
- Pré-validado com a plataforma NVIDIA AI Workbench pronta para desenvolvedores
- Caso de uso inicial do chatbot implementado com seus dados
- Abordagem econômica e de baixo risco para experimentar e desenvolver habilidades de IA generativa

WORKSTATION MÓVEL DELL PRECISION 5690/7780 COM GPUS NVIDIA

COMECE HOJE MESMO

DELL Technologies AI Factory WITH NVIDIA