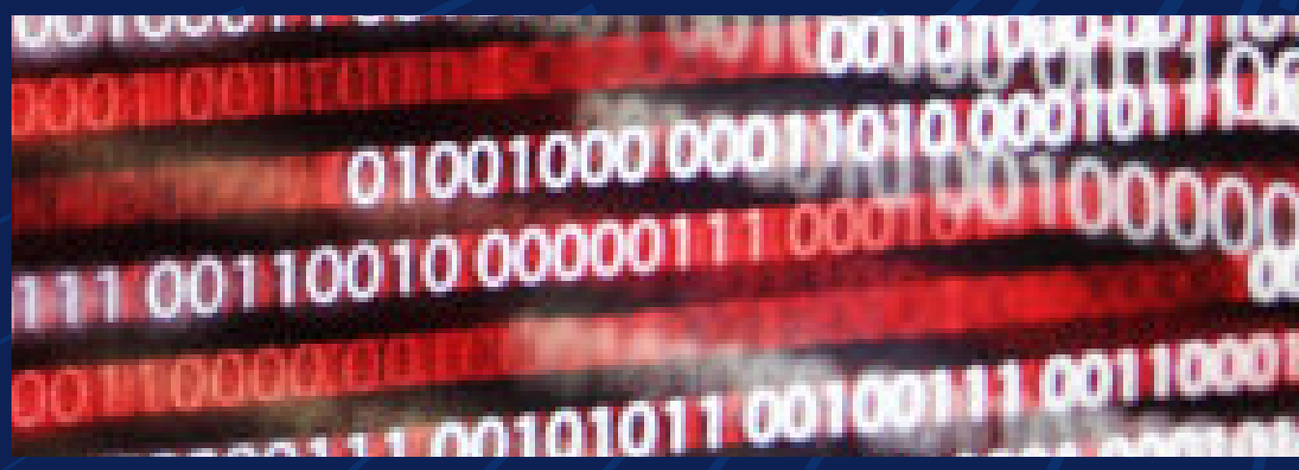
Desmistificando a segurança cibernética:

Desmascarando mitos sobre segurança de IA



A IA está transformando as indústrias, mas quando se trata de proteger a IA, muitas organizações são vítimas de mitos que fazem com que ela pareça mais complexa do que realmente é. A verdade? Proteger sistemas de IA não exige começar completamente do zero, aplicar os princípios de segurança cibernética existentes aos desafios exclusivos da IA é uma grande ajuda.

Na Dell Technologies, entendemos a arquitetura por trás da IA e podemos ajudar você a adaptar suas soluções atuais para se adequar a essa nova estrutura. Vamos desmistificar os mitos mais comuns em torno da segurança da IA e descobrir as verdades para ajudar você a proteger seus sistemas de forma eficaz.

Mito 1: "Os sistemas de IA são muito complexos para serem protegidos".

A verdade: É verdade que a IA cria novos riscos de segurança cibernética, como injeção de prompts, manuseio de dados e divulgação de informações confidenciais, só para citar alguns. Além disso, os sistemas de IA agêntica também apresentam uma superfície mais ampla de ataque, pois podem ser explorados para manipular resultados ou aumentar privilégios.

Dito isso, embora seja essencial reconhecer essas vulnerabilidades e implementar medidas de segurança para proteger os sistemas de IA contra ameaças tradicionais e específicas da IA, os riscos podem ser gerenciados e os modelos de IA podem ser protegidos. É importante ter em mente que os sistemas de IA exigem quantidades significativas de dados como entradas e criam grandes quantidades de dados como saídas. Isso coloca a proteção de dados no centro das estratégias de segurança, juntamente com:

- Princípios de Zero Trust, como gerenciamento de identidades, acesso baseado em funções e verificação contínua.
- Testes de violação regulares e gerenciamento de vulnerabilidades para identificar pontos fracos.
- Registro e auditoria para validar entradas e saídas de dados

Mito 2: "Nenhuma das minhas ferramentas existentes protegerá a IA."

A verdade: A segurança da IA não significa começar do zero, mas sim usar de forma mais inteligente as ferramentas que você já tem. A maioria das ferramentas de segurança cibernética existentes pode ser adaptada para proteger os sistemas de IA de forma eficaz. Em sua essência, a IA é mais uma carga de trabalho impulsionando seus negócios, ainda que com características únicas. Práticas básicas de segurança cibernética, como gerenciamento de identidades, segmentação e monitoramento de rede, proteção de endpoints e proteção de dados, permanecem essenciais para proteger ambientes de IA. O essencial é adaptar essas práticas para enfrentar desafios específicos da IA, como proteger os dados de treinamento, garantir a segurança dos algoritmos e mitigar riscos como entradas de adversários.

Uma defesa forte começa com uma boa higiene cibernética, como aplicação de patches no sistema, controle de acesso e gerenciamento de vulnerabilidades. O importante é adaptar essas práticas para abordar riscos específicos da IA. Com estratégias focadas em IA integradas à sua abordagem de segurança atual e as ferramentas certas, a segurança da IA se torna gerenciável e eficaz.

No entanto, é importante mencionar que o hardware atualizado pode desempenhar um papel fundamental no combate a ataques cibernéticos. Por exemplo, os AI PCs modernos criam uma primeira linha de defesa sólida contra um dos principais vetores de ataque: os endpoints. Com o fim do suporte do Windows 10, os PCs desatualizados se tornam um risco. Além disso, o Windows 11 requer o Trusted Platform Module (TPM) versão 2.0, um chip de segurança que ajuda com a criptografia, inicialização segura e proteção contra ataques de firmware. Muitos PCs mais antigos não têm TPM ou só oferecem suporte a uma versão mais antiga. A Dell oferece AI PCs comerciais seguros com esses aprimoramentos incorporados.

O mesmo vale para infraestrutura de IA, como servidores e armazenamento. A Dell AI Factory inclui hardware otimizado para segurança de IA e contém uma série de recursos de segurança integrados, desde uma cadeia de suprimentos segura até imutabilidade de dados para isolamento e criptografia.

Mito 3: "A segurança da IA se limita à proteção de dados."

A verdade: A segurança da IA vai além da proteção básica de dados, ela envolve a proteção de todo o ecossistema da IA, incluindo modelos, APIs, resultados, sistemas e dispositivos. À medida que a IA se torna mais integrada a aplicativos essenciais, os riscos associados ao seu uso indevido ou exploração aumentam. Sem medidas de segurança robustas, os modelos de IA podem ser adulterados para gerar resultados prejudiciais ou enganosos, as APIs podem ser exploradas para obter acesso não autorizado a sistemas confidenciais e os resultados podem expor inadvertidamente informações privadas ou confidenciais.

A segurança abrangente da IA requer uma abordagem em várias camadas. Isso inclui a proteção de modelos contra ataques de adversários que tentam manipular dados de entrada para enganar sistemas de IA, proteger APIs com métodos de autenticação fortes para impedir o uso não autorizado e **monitorar continuamente os resultados** em busca de padrões incomuns ou suspeitos que possam sinalizar um ataque ou mau funcionamento. A segurança eficaz da IA não apenas garante a integridade e a confiabilidade dos sistemas de IA, mas também gera confiança com usuários e partes interessadas, reduzindo os riscos de uso mal-intencionado ou consequências não intencionais.

Mito 4: "A IA não precisa de supervisão humana".

A verdade: A governança e a supervisão humana são fundamentais para garantir que os sistemas de IA operem de forma ética, previsível e alinhada aos valores humanos. Os sistemas avançados de IA, particularmente a IA agêntica com capacidade de decisão autônoma, apresentam desafios únicos que exigem proteções robustas. Sem supervisão adequada, esses sistemas podem se desviar dos objetivos pretendidos ou apresentar comportamentos não intencionais que podem representar riscos.

Para resolver isso, é essencial estabelecer limites claros, implementar mecanismos de controle em camadas e garantir o envolvimento humano contínuo em processos críticos de tomada de decisão. Auditorias regulares, transparência nas operações de IA e testes completos podem aumentar ainda mais a responsabilização e a confiança, ajudando a prevenir o uso indevido e promovendo a implementação responsável das tecnologias de IA.

Melhores práticas para fortalecer a segurança da IA

Para preencher lacunas de segurança específicas da IA, as organizações precisam adotar uma abordagem proativa e estratégica. Aqui estão 10 melhores práticas para proteger seus sistemas de IA:

Arquitetura de segurança em camadas:
Use segmentação, firewalls e autenticação forte para proteger sua infraestrutura, software e dados em todas as camadas.

Proteção da cadeia de suprimentos:
Implemente um forte programa de gerenciamento de fornecedores. Audite fornecedores e componentes de terceiros, valide a integridade e confie em códigos assinados para prevenir vulnerabilidades no ciclo de desenvolvimento de IA.

Proteger dados e modelos de treinamento:
Proteja-se contra dados comprometidos, entradas de adversários e outras ameaças, monitorando a integridade dos dados e aplicando ferramentas de validação robustas.

Reforçar controles de acesso:
Aplique os princípios de privilégios mínimos, implemente o controle de acesso baseado em função (RBAC), troque as credenciais regularmente e audite permissões para impedir o acesso não autorizado.

APIs seguras:
Use protocolos de autenticação sólidos (como OAuth 2.0), use criptografia HTTPS e atualize APIs regularmente para encerrar possíveis vulnerabilidades.

Monitorar e validar resultados da IA:
Use detecção de anomalias, registros e alertas para monitorar padrões incomuns ou comportamentos prejudiciais nos resultados da IA.

Planeje a resiliência:
Faça backup regularmente dos dados e teste planos de recuperação de desastres para minimizar o tempo de inatividade e garantir uma recuperação rápida em caso de violação.

Implementar uma criptografia robusta:
Criptografe dados confidenciais em repouso e em trânsito utilizando algoritmos robustos e gerencie e gire as chaves de criptografia de forma segura e regular.

Realize auditorias regulares de segurança e testes de violação:
Avalie frequentemente os sistemas em busca de vulnerabilidades e use testes de violação para descobrir riscos antes que eles possam ser explorados.

Treinamento da equipe sobre as melhores práticas de segurança de IA:
Treine regularmente sua equipe sobre desenvolvimento seguro, reconhecimento de ameaças e manutenção de práticas sólidas de segurança para evitar violações.

DELL Technologies

Proposta de valor da Dell: Soluções práticas de segurança de IA.

A segurança da IA pode parecer complexa, mas não é tão assustadora quanto parece. A verdade? Proteger a IA não é tão diferente de proteger suas cargas de trabalho existentes, trata-se de entender a arquitetura e aplicar as estratégias certas. É aí que entra a Dell Technologies.

Desmistificamos a segurança da IA aproveitando suas soluções atuais e integrando-as perfeitamente em arquiteturas focadas em IA. Enfrentamos desafios como injeção de prompts, abuso de API e ataques de adversários sem exigir uma revisão completa da infraestrutura.

A experiência da Dell está em derrubar os mitos sobre a segurança de IA e demonstrar sua real viabilidade. Não importa se você está apenas começando sua jornada de IA ou quer aprimorar suas defesas, nós ajudaremos você a proteger seus investimentos, proteger seus sistemas e construir um futuro digital resiliente, com confiança e eficácia. Vamos simplificar a segurança da IA juntos.

Produtos e soluções da Dell que podem ajudar

Solução Dell em destaque	Descrição
Dell AI Factory	A Dell AI Factory protege as cargas de trabalho de IA por meio de uma cadeia de suprimentos segura, garantindo uma infraestrutura confiável do desenvolvimento à implantação. Com recursos como imutabilidade, isolamento e criptografia de dados, ela protege modelos e conjuntos de dados confidenciais, defende contra ameaças cibernéticas e permite operações de IA escaláveis, eficientes e contínuas em ambientes dinâmicos orientados por dados.
Resiliência cibernética	O PowerProtect protege cargas de trabalho de IA com recursos avançados, como imutabilidade e isolamento, garantindo a integridade e a proteção dos dados contra ameaças cibernéticas. Ele oferece criptografia completa e detecção de anomalias, além de permitir a recuperação rápida para minimizar o tempo de inatividade.
Dell Trusted Workspace (Segurança de endpoints)	Uma combinação de recursos nativos e opcionais adicionais projetados para proteger PCs comerciais de IA e as cargas de trabalho de IA em execução neles. Desenvolvida com práticas seguras na cadeia de suprimentos, os recursos integrados incluem SafeBIOS e SafeID com TPM. Os complementos opcionais incluem Secured Component Verification, SafeID com ControlVault e software de parceiros como CrowdStrike e Absolute para maximizar a segurança do espaço de trabalho.
Serviços de consultoria em segurança de IA	Um conjunto de serviços que pode ajudar você a desenvolver e implementar uma estratégia abrangente de segurança de IA. As ofertas incluem serviços de consultoria, vCISO de IA e planejamento de segurança de dados.
Operações de segurança gerenciadas para IA	Permite uma visibilidade profunda em toda a pilha para detectar e responder rapidamente às ameaças. Os recursos incluem Managed Detection and Response, proteção de IA gerenciada, testes de violação para IA e serviços de resposta e recuperação de incidentes.
Integração do software de segurança	Projete, instale e configure ferramentas de segurança que protejam o gerenciamento de acesso, aplicativos, redes, nuvens e muito mais.

Saiba como enfrentar alguns dos principais desafios atuais de segurança cibernética em **dell.com/cybersecuritymonth**