netskope

Protegendo a IA: 5 conversas cruciais para CISOs

Conteúdo

Introdução: Um mandato duplo para a IA

A crescente importância da tecnologia nas organizações modernas de hoje colocou o departamento de TI em evidência como nunca antes. Como resultado, os líderes de TI se veem envolvidos em uma série de conversas cruciais com seus CEOs, conselhos e pares da alta liderança, buscando estabelecer as melhores maneiras pelas quais a tecnologia pode contribuir para o sucesso dos negócios — e nenhum tópico é mais proeminente nessas discussões do que a IA.

A IA representa um dilema particular para os CIOs, CISOs e suas equipes. No relatório Crucial Conversations da Netskope[1], revelamos que os CEOs estão dando aos seus líderes de TI uma faca de dois gumes: integrar a IA para incentivar a experimentação e gerar valor de negócio mensurável, mas também reduzir custos, atuar como guardiões contra gastos excessivos, evitar exageros e proteger-se contra possíveis vazamentos de dados ou violações de segurança.

Resumindo, os líderes de TI devem usar a IA para viabilizar inovações disruptivas, ao mesmo tempo que se protegem dos riscos que ela acarreta. É uma dualidade que coloca uma pressão significativa sobre os ombros deles.

Cada organização encontra-se em um estágio diferente de maturidade em IA. Algumas empresas ainda estão identificando os casos de uso em que a IA pode causar impacto, enquanto outras estão avançando rapidamente, criando seus próprios aplicativos de IA e incentivando os funcionários a adotarem amplamente essas ferramentas. Todos estão correndo para utilizar a IA para acelerar seu crescimento, mas estão se movendo em velocidades diferentes, partindo de pontos de partida diferentes.

[1] https://www.netskope.com/crucial-conversations

Cinco etapas para a adoção da IA

Independentemente do estágio de maturidade em que sua empresa se encontre em relação à IA, existem considerações de segurança vitais que precisam ser levadas em conta. A chave é planejar sua estratégia de segurança com uma compreensão completa dos riscos em cada etapa.

1. **Podemos experimentar ferramentas de IA** e ao mesmo tempo gerenciar os riscos de shadow AI?
2. **Podemos utilizar a IA incorporada em plataformas SaaS** sem permitir o compartilhamento de dados não aprovados?
3. **Como gerenciamos aplicativos de IA independentes** e evitamos vazamento de dados?
4. Como podemos evitar respostas de modelo prejudiciais ou tendenciosas e vulnerabilidades comuns de aplicativos ao **construir aplicativos de IA privados?**
5. Como evitamos conceder permissões excessivas ao **implantar agentes autônomos**?

Experimentação

IA integrada em plataformas SaaS

Aplicativos de IA independentes autorizados

Aplicativos de IA privados

Agentes autônomos

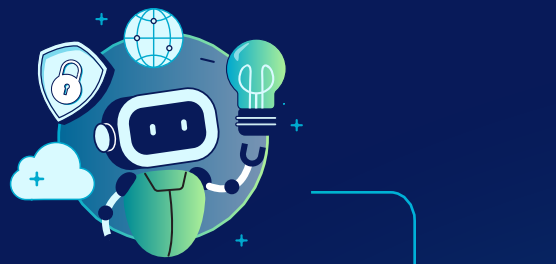
Rumo a conversas mais produtivas com o alto escalão

A inteligência artificial não é apenas o principal assunto de conversa nos círculos tecnológicos hoje em dia, mas também uma das principais prioridades entre os executivos de alto escalão e os conselhos de administração. Com base em nossa pesquisa com CEOs [1], sabemos que eles estão entusiasmados com o potencial da IA e querem que seus líderes de TI criem um caminho para a adoção e integração, quando apropriado, sem se deixarem levar pela propaganda do setor.

O desafio para os profissionais de TI, especialmente na área de segurança, é implantar IA de uma forma que evite sacrifícios entre desempenho e segurança, e que permaneça em conformidade, reduzindo custos e complexidade. Da mesma forma, à medida que os planos de implementação se tornam mais detalhados, é fundamental que os benefícios e os riscos para o negócio permaneçam em vista.

Com este e-book, esperamos contribuir para esses resultados. ***Protegendo a IA: 5 conversas cruciais para CISOs*** tem como objetivo ajudar as equipes de segurança a avançarem com confiança em suas jornadas de IA, mantendo conversas mais produtivas sobre os desafios e oportunidades da IA com os colegas. O objetivo final é ajudar as organizações a incorporar princípios de IA em sua estratégia de segurança e princípios de segurança em sua estratégia de IA – transformando a segurança em um motor de crescimento, em vez de um obstáculo.

[1] *https://www.netskope.com/crucial-conversations*



Inovar com IA, defender-se contra riscos — três princípios a priorizar.

Independentemente do estágio em que você esteja em sua jornada de adoção de IA, três princípios se aplicam para aproveitar com segurança o potencial da tecnologia.

1. **Visibilidade.** As equipes de segurança precisam de visibilidade total sobre seu cenário de IA, para que saibam quais ferramentas de IA estão sendo usadas e como.

2. **Proteção.** Os profissionais precisam implementar proteção contextual para que a segurança dinâmica e adaptável proteja os negócios sem impedir a inovação.

3. **Prontidão.** Ao compreenderem proativamente os dados e aplicativos, os profissionais de segurança podem alcançar a prontidão em IA que prepara a sua organização para o sucesso.

1ª Etapa: Experimentação

Podemos experimentar ferramentas de IA e ao mesmo tempo gerenciar os riscos de shadow AI?

Quando o ChatGPT foi lançado em novembro de 2022, ele conquistou o mundo — e pegou as empresas de surpresa. Quase imediatamente, os funcionários começaram a usar instâncias pessoais de chatbots de IA para ajudar a agilizar ou resolver suas tarefas de trabalho. Atualmente, a shadow AI continua sendo um problema constante para muitas organizações: De acordo com a pesquisa do Netskope Threat Labs[1], em 2025, um número impressionante de 72% dos usuários corporativos ainda utilizavam contas pessoais para acessar o ChatGPT, o Google Gemini e outros aplicativos populares de IA generativa no trabalho.

Essa questão está ficando cada vez mais complicada. Quase todos os aplicativos SaaS consolidados incluem funcionalidades de IA integradas, os modelos de IA agora se comunicam diretamente entre si, os agentes podem ser criados por meio de linguagem natural, deixando de ser exclusividade de especialistas técnicos, e todas essas instâncias de IA interagem com mais dados e aplicativos do que um ser humano jamais conseguiria. Como resultado, a shadow AI está se expandindo a uma taxa sem precedentes.

72% dos usuários corporativos ainda utilizam contas pessoais para acessar os aplicativos de IA generativa no trabalho.

Netskope, Generative AI Cloud and Threat Report 2025

[1] https://www.netskope.com/resources/reports-guides/cloud-and-threat-report-generative-ai-2025

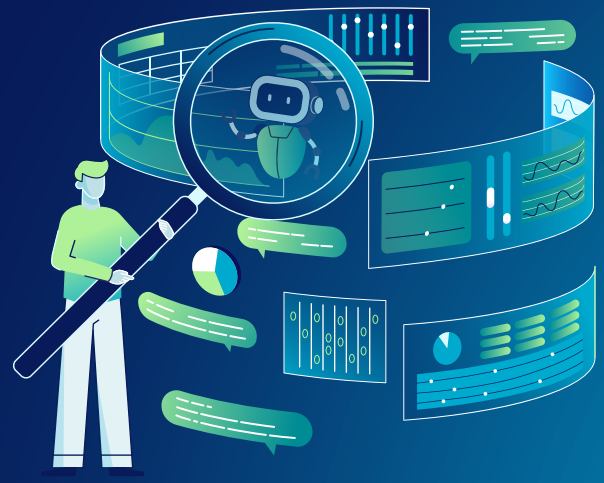
As equipes de segurança precisam urgentemente de uma visão ampla e aprofundada do cenário de IA em que atuam. Elas devem garantir que tenham visibilidade de toda a gama de ferramentas de IA utilizadas em sua organização, incluindo aplicativos não gerenciados e instâncias pessoais. Elas também precisam se aprofundar e entender o que os usuários e agentes estão fazendo nessas interações.

Somente esse nível de conhecimento aprofundado permitirá que as equipes de segurança ultrapassem a confiança cega e obtenham controle estratégico sobre as atividades de IA de suas organizações.

A dura realidade é que você não pode proteger o que não pode ver.

Como fazemos isso

Por meio do Netskope One Cloud Access Security Broker (CASB) e do Next Generation Secure Web Gateway (NG-SWG), as empresas podem obter visibilidade detalhada das atividades de IA em seu ambiente. Nosso painel de IA fornece detalhes minuciosos sobre quais usuários estão acessando quais aplicativos, bem como quais ações estão realizando. Também auxilia as empresas a realizar a descoberta e inspeção inline de todas as interações com LLMs públicos (dados em trânsito), incluindo as interações entre usuários e aplicativos.



2ª Etapa: Inteligência artificial integrada em plataformas SaaS

É possível aproveitar a IA integrada em plataformas SaaS sem permitir o compartilhamento de dados não autorizados?

Os LLMs e os aplicativos de IA dedicados não são mais os únicos vetores de risco de uma perspectiva de IA. Com a evolução da tecnologia, os recursos de IA estão sendo incorporados a um número cada vez maior de aplicativos SaaS — desde plataformas de videochamadas até ferramentas de produtividade e sistemas de gestão de vendas.

Essas ferramentas SaaS geralmente estão profundamente integradas às empresas modernas, que dependem delas para funções essenciais do dia a dia, tornando quase impossível bloqueá-las ou removê-las. E os recursos de IA normalmente aceleram a produtividade de uma forma que nenhuma organização gostaria de impedir.

A nova funcionalidade de IA costuma ser adicionada com o mínimo de obstáculos — por exemplo, incluída em uma atualização geral com pouquíssimas informações sobre os termos e condições de uso dos dados. Um aplicativo de videochamadas, por exemplo, poderia ativar por padrão um recurso de IA para anotações, que registraria e armazenaria informações confidenciais da empresa. Isso pode facilmente pegar as equipes de segurança desprevenidas.

Um aplicativo SaaS existente pode introduzir novas funcionalidades de IA — e até mesmo ativá-las automaticamente — potencialmente pegando as equipes de segurança desprevenidas.

Os profissionais de segurança precisam estar a par do cenário de aplicativos SaaS, com visibilidade dos recursos de IA, suas funções e os termos e condições contratuais associados à governança de dados. Isso deve incluir a compreensão de fatores como a forma como cada aplicativo usa IA, se ele usa seus dados para treinar seus modelos, se está em conformidade com as principais regulamentações e se seus recursos de IA podem ser desativados.

As organizações também devem considerar seriamente a categorização e classificação de dados sensíveis para poderem aplicar políticas específicas que protejam a propriedade intelectual da empresa ou dados regulamentados (por exemplo), ao mesmo tempo que permitem regras mais flexíveis em relação a informações não sensíveis.

Como fazemos isso

O Cloud Confidence Index (CCI) da Netskope é um repositório com mais de 85.000 aplicativos SaaS, que fornece um contexto de risco abrangente e permite que as equipes de segurança tomem decisões informadas sobre quais aplicativos com IA permitir, restringir ou bloquear.

3ª Etapa: Aplicativos de IA independentes gerenciados

Como gerenciamos aplicativos de IA independentes e evitamos vazamento de dados?

Atualmente, muitas organizações já escolheram sua ferramenta de IA preferida, como o ChatGPT da OpenAI, o Copilot da Microsoft, o Gemini da Google ou o Claude da Anthropic. A padronização em torno de um único sistema em toda a empresa traz vantagens óbvias em termos de recursos prontos para uso corporativo, aprendizado reforçado e proteções de segurança. E se a organização também bloquear outros sistemas de IA, isso reduz ainda mais a superfície de ataque potencial.

No entanto, essa abordagem não elimina completamente o risco. Uma ferramenta de IA corporativa só se torna verdadeiramente valiosa se estiver integrada a outros documentos e fontes de informação dentro da empresa. Isso ainda poderia permitir que usuários individuais extraíssem dados de documentos internos aos quais não deveriam ter acesso — em outras palavras, causando vazamento de dados dentro da sua organização.

Um funcionário do departamento de marketing poderia perguntar a uma IA empresarial com permissões excessivas sobre os próximos recursos no roteiro do produto e receber informações extraídas de documentos confidenciais aos quais não deveria ter acesso.

Como fazemos isso

A Netskope protege ativamente contra ameaças específicas de IA durante o runtime. Caso um usuário tente inserir informações confidenciais, o Netskope One Data Loss Prevention (DLP) intervém instantaneamente, bloqueando a entrada de informações de identificação pessoal, código-fonte ou segredos proprietários no modelo de IA. Isso também pode acionar uma janela pop-up de orientação para instruir o usuário.

Simultaneamente, o Netskope One AI Guardrails fornece moderação de conteúdo em tempo real para cada interação. Ele analisa a intenção por trás de prompts e respostas para bloquear automaticamente ataques maliciosos sofisticados, como injeções de prompt e tentativas de jailbreak. Além disso, o Guardrails promove o uso responsável da IA, filtrando conteúdo prejudicial ou discriminatório e bloqueando a distribuição de materiais protegidos por direitos autorais. Ao combinar essas funcionalidades do DLP e do Guardrail, as organizações podem orientar proativamente os usuários, protegendo todo o ecossistema de IA contra vazamentos de dados e ameaças emergentes.

4ª Etapa: Aplicativos de IA privados

Como podemos evitar respostas de modelo prejudiciais ou tendenciosas e vulnerabilidades comuns de aplicativos ao construir aplicativos de IA privados?

Organizações em setores altamente regulamentados (como saúde, serviços financeiros e governo) estão na vanguarda do desenvolvimento de aplicativos privados de IA. À medida que a confiança na IA aumenta, muitos estão recorrendo a modelos executados localmente e treinados com os próprios dados da organização para reduzir os riscos relacionados à residência de dados, privacidade, conformidade e exposição a terceiros, ao mesmo tempo que melhoram a relevância e a confiabilidade.

Um terço das organizações já utiliza os serviços da OpenAI via Azure, 27% utilizam o Amazon Bedrock e 10% estão desenvolvendo soluções com base no Google Vertex AI[1]. Todas essas plataformas de nível empresarial oferecem serviços de IA seguros e baseados em nuvem que fornecem controles de privacidade mais fortes e opções de integração mais profundas do que suas versões públicas.

No entanto, a construção de IA privada também transfere a responsabilidade pela segurança para a organização. Além da proteção em runtime contra ameaças específicas de IA e uso indevido por funcionários, uma superfície de ataque adicional reside nas ferramentas usadas para projetar e implantar esses sistemas, que podem não possuir proteções integradas.

Um aspecto importante a considerar ao operar um modelo de IA localmente é se ele é suscetível a vulnerabilidades. Se uma organização personalizar um modelo de código aberto, por exemplo, a equipe de segurança ainda deverá testar rigorosamente o código e garantir que nenhum componente malicioso tenha sido introduzido — por exemplo, código que possa capturar ou transmitir prompts para uma fonte externa.

Outro ponto a considerar é confirmar se os dados de treinamento da organização não contêm informações indesejadas. As equipes devem verificar se há algum elemento tendencioso, sensível ou prejudicial nesses conjuntos de dados, que geralmente são muito grandes.

Como fazemos isso

Ao centralizar a autenticação, o gerenciamento de tráfego e a inspeção de conteúdo entre aplicativos privados e LLMs, o Netskope One AI Gateway garante que os fluxos de dados autônomos permaneçam controlados e seguros. Além disso, o Netskope One AI Red Teaming realiza testes de estresse proativos em modelos personalizados, automatizando simulações adversárias em pipelines CI/CD para descobrir vulnerabilidades como injeções de prompt.

O Netskope One AI Guardrails mitiga ataques sofisticados — incluindo tentativas de injeção rápida e jailbreak — por meio da análise em tempo real de todo o tráfego, além de funcionar como um moderador de conteúdo para identificar e controlar conteúdo prejudicial ou discriminatório, tanto para interações humanas quanto para interações com agentes.

Além disso, com o Netskope One DSPM, os líderes de segurança podem obter visibilidade e controle sobre seus dados, onde quer que estejam. Isso os ajuda a descobrir e classificar informações sensíveis, por exemplo, que podem ser usadas para treinar um modelo de IA.

[1] *Netskope Threat Labs, Netskope Cloud and Threat Report 2026*

5ª Etapa: Agentes autônomos

Como evitamos conceder permissões excessivas ao implantar agentes autônomos?

A IA agêntica é a mais recente queridinha da onda de inteligência artificial. De fato, muitos analistas a apontam como parte fundamental do futuro da tecnologia corporativa, e a empresa de análises Gartner® prevê que, até 2028, pelo menos 15% das decisões diárias de negócios serão tomadas de forma autônoma por meio de IA agêntica, partindo de 0% em 2024[1].

Embora a implementação dessa tecnologia ainda esteja em seus estágios iniciais, uma pesquisa do Netskope Threat Labs, de agosto de 2025, descobriu uma massa crítica de usuários em organizações que já estão criando agentes de IA ou utilizando os recursos de IA agêntica em soluções SaaS.

Por exemplo, o GitHub Copilot agora é usado em 39% das organizações, e 5,5% têm usuários executando agentes gerados a partir de estruturas populares de agentes de IA em infraestruturas locais. Segundo os pesquisadores, 66% das organizações têm usuários fazendo chamadas de API para api.openai.com e 13% para api.anthropic.com[2].

39% das organizações usam o GitHub Copilot e 5,5% executam agentes de IA gerados a partir de infraestruturas populares locais.

Netskope, Cloud and Threat Report: Shadow AI and Agentic AI 2025

[1] *Press release da Gartner: Gartner identifica as 10 principais tendências estratégicas de tecnologia para 2025, 21 de outubro de 2024*

[2] *https://www.netskope.com/resources/cloud-and-threat-reports/cloud-and-threat-report-shadow-ai-and-agentic-ai-2025*

Atualmente, muitas empresas não têm uma noção clara da extensão de seu ambiente de IA agêntica. Com este campo evoluindo tão rapidamente e com novas funcionalidades sendo adicionadas o tempo todo, a shadow AI agêntica é uma faceta cada vez mais significativa do problema geral da shadow AI.

À medida que a adoção de agentes de IA cresce — juntamente com a amplitude de suas capacidades em toda a organização — os riscos de segurança que os acompanham se multiplicarão. Será imprescindível que as equipes compreendam as ações realizadas por cada agente e implementem controles e políticas adequados para gerenciar as permissões e atividades executadas.

Os aplicativos com inteligência artificial dependem da comunicação autenticada entre aplicativos internos, agentes autônomos e LLMs hospedados de forma privada. Para isso, eles usam o MCP (model context protocol) e APIs, mas, embora os protocolos em si sejam rotas de comunicação seguras, essas interações não humanas abrem uma brecha crítica de segurança. APIs e MCP permitem que agentes de IA interajam diretamente com dados e ferramentas sensíveis, contornando a segurança tradicional centrada no humano. Essa lacuna cria o risco de interações autônomas sem supervisão, levando a vazamentos de credenciais, envenenamento de ferramentas maliciosas e exfiltração de dados não autorizada.

Como fazemos isso

O Netskope One AI Gateway funciona como um gateway definido por software para interceptar e controlar o tráfego de API entre aplicativos internos, agentes autônomos e LLMs hospedados privadamente, garantindo que apenas agentes autenticados possam se comunicar com os LLMs, exigindo um token válido gerado pelo AI Gateway para cada solicitação.

O Netskope One Agentic Broker oferece visibilidade unificada e proteção em tempo real para aplicativos habilitados para MCP — incluindo editores de código de IA, interfaces de bate-papo e ferramentas de desenvolvedor — decodificando e protegendo o tráfego MCP entre agentes de IA e fontes de dados: preenchendo a lacuna entre as interações humano-LLM e os fluxos de trabalho de IA de máquina para máquina. Isso garante uma postura de segurança consistente que protege dados corporativos sensíveis, ao mesmo tempo que permite a velocidade e a escalabilidade da automação agêntica.

Conclusão: Gerenciando riscos de IA sem sacrifícios

A IA representa um desafio e uma oportunidade que definirão uma era para os CIOs, CISOs e suas equipes. Isso os vincula mais estreitamente à estratégia e ao crescimento dos negócios do que nunca, ampliando sua influência e impacto. Mas também representa riscos significativos e em rápida evolução para os dados, receitas e reputação da organização — aumentando a gravidade de cada violação e invasão.

Para navegar nesse cenário, os líderes de TI precisam de um modelo claro para entender as possibilidades disruptivas da IA e os requisitos de defesa — o que lhes dá confiança para falar sobre IA com colegas não técnicos, de modo a gerenciar os riscos potenciais da IA e atender a rigorosos padrões de conformidade.

Para os profissionais de TI e segurança de hoje, a chave é garantir a adoção da IA de ponta a ponta para promover a inovação segura, mantendo ao mesmo tempo a resiliência das operações comerciais.

Netskope One é uma plataforma única e consolidada que oferece uma maneira de gerenciar os riscos da IA sem sacrificar o desempenho ou a experiência do usuário, reduzindo a complexidade e garantindo a conformidade simultaneamente.

À medida que mais empresas avançam em sua jornada de adoção de IA — da experimentação inicial à implantação automatizada — os líderes de TI podem desempenhar um papel fundamental na promoção e viabilização da inovação nos negócios. Ao explorar com segurança os benefícios da IA, os CIOs e CISOs de hoje podem gerar um impacto nos negócios que leve suas organizações a um novo patamar.

O **Netskope One AI Security** oferece uma solução única para governar seu ecossistema de IA e proteger seus dados. Ele protege usuários e agentes automatizados em SaaS público, ferramentas de IA privadas e fluxos de trabalho baseados em agentes. Combinando alto desempenho com controles zero trust baseados em contexto, a Netskope permite que as organizações passem da experimentação com IA para a obtenção de vantagens competitivas em IA.

Leia mais sobre o que os CEOs esperam de seus líderes de TI no relatório Crucial Conversations da Netskope aqui.

Sobre a Netskope

A Netskope é líder em segurança moderna, redes e analytics para a era da nuvem e da IA. A arquitetura exclusiva da plataforma Netskope One permite segurança em tempo real e baseada no contexto para pessoas, dispositivos e dados onde quer que estejam, além de otimizar o desempenho da rede — sem concessões ou sacrifícios. Milhares de clientes e parceiros confiam na plataforma Netskope One, em seu Zero Trust Engine patenteado e em sua poderosa rede NewEdge para reduzir riscos, simplificar a infraestrutura convergente e fornecer visibilidade e controle completos sobre a atividade em nuvem, IA, SaaS, web e aplicativos privados.

Interessado em saber mais?

Peça uma demonstração

netskope

04/26 EB-950-2-PT