Quando a IA se torna uma arma cibernética

Deepfakes, ataques de phishing automatizados e engenharia social impulsionada por IA estão em ascensão.

Descubra como proteger sua empresa.

Penso

Apoio: veeam Adistec

Introdução

Quando a IA se Torna uma Arma Cibernética

A Inteligência Artificial (IA) está cada vez mais presente em nossas vidas, permeando tanto os aspectos produtivos quanto os criminosos. Neste eBook, baseado no webinar exclusivo realizado pela Penso em 14 de maio de 2025, exploraremos como a IA generativa está sendo utilizada como uma arma cibernética, os riscos que ela representa e, principalmente, como as empresas podem se proteger. Apresentado por Thiago Madeira de Lima, CEO da Penso, este conteúdo visa gerar insights valiosos para o cenário atual de cibersegurança.

Deepfakes, Phishing e Engenharia Social Assistida por IA.

Saiba como se proteger

Apoio: veeam Adistec

"SEM PÂNICO, SARAH CONNOR.

A PENSO ESTÁ AQUI

PARA GARANTIR QUE A IA SEJA SUA ALIADA, NÃO SUA ADVERSÁRIA."

Sobre a **Penso**

A Penso é parceira estratégica da **VEEAM**, líder global em backup e recuperação.

Com as mais altas certificações do mercado e **ISO-27001**, oferecemos **Backup em Nuvem, Backup Imutável e DRaaS**, garantindo segurança máxima e recuperação rápida para seus dados.

+22 Anos de know-how tecnológico	**+1600** Clientes de grande porte	**+250** Especialistas em tecnologia	**+5** Data Centers Tier III no Brasil	**98%** de satisfação com nossas soluções

Acesse nosso site e conheça as **soluções da Penso**

Protegemos seus dados para que sua empresa nunca pare.

- Linhas de defesa contra ransomware
- Disaster Recovery as a Service (DRaaS)
- Cyber Security
- BaaS para Microsoft 365
- Backup Cloud e replicação

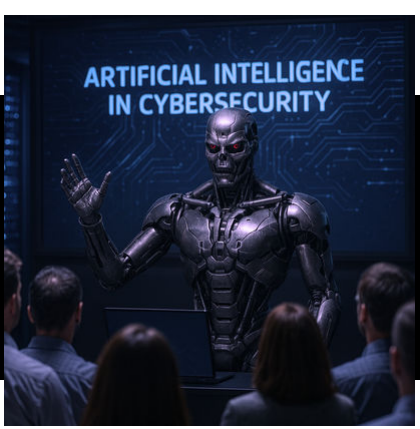
veeam
Competency Partner
DRaaS
BaaS *for 365*
Off-site Backup

Com mais de 22 anos de experiência, a **Penso é uma referência em soluções tecnológicas, reconhecida como Impact Cloud and Service Provider Partner of the Year Brasil pela Veeam.**

Esse prêmio destaca nosso papel como parceiro de maior impacto no ano, além de sermos uma das empresas mais certificadas entre os parceiros Veeam.

Com uma equipe de **mais de 250 profissionais,** oferecemos backup em nuvem, **backup imutável e DRaaS, garantindo segurança e recuperação rápida para grandes empresas.**

ÍNDICE



SPEAKER

Thiago Lima

CEO na Penso

+ 1.600 clientes corporativos ativos

+ Especialista em resiliência de dados

"Acredito que dados resilientes e sistemas disponíveis são a base de qualquer negócio moderno. **É isso que entregamos todos os dias**."

Apoio: veeam Adistec

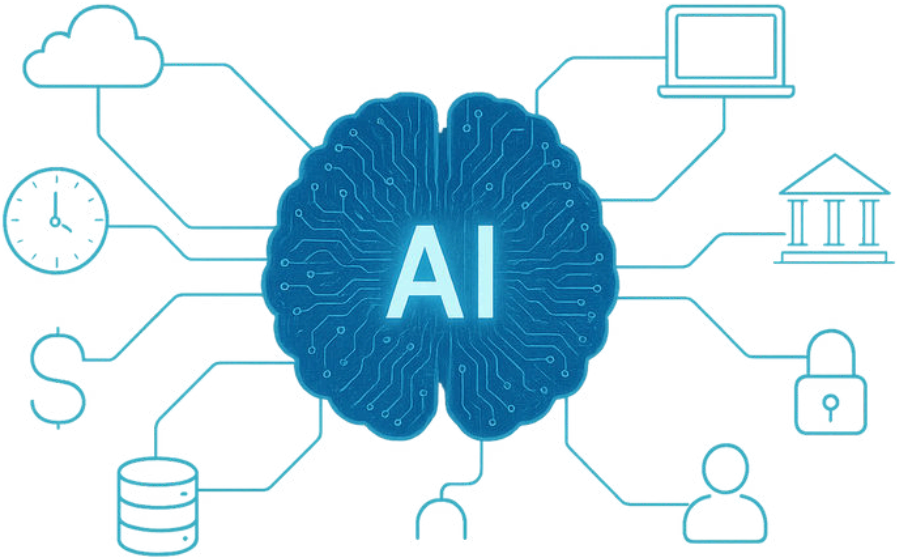
Você é capaz de dizer se essas pessoas são reais ou Deep Fake?

É por isso que esse conteúdo é tão necessário

....

A Inteligência Artificial está presente em todos os lugares

Inclusive no Cibercrime



Adoção de IA Generativa cresceu de 31% para 71% entre 2023 e 2024

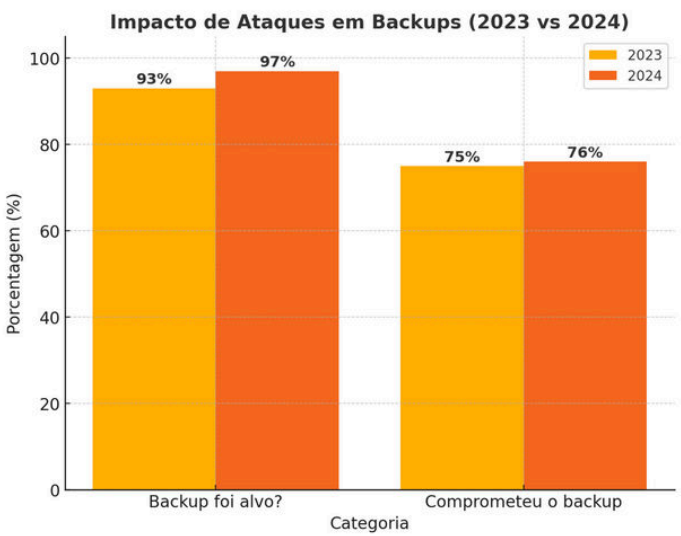
Capítulo 1: O Crescimento da IA Generativa e Seus Riscos

A Adoção da IA Generativa nas Empresas

Entre 2023 e 2024, a adoção da IA generativa nas empresas cresceu de 31% para 71%, um aumento de quase 250%. Ferramentas como ChatGPT e Copilot impulsionam essa expansão, trazendo benefícios significativos para a produtividade. No entanto, há também um lado sombrio: o uso malicioso da IA está aumentando ainda mais rapidamente.

O Impacto da IA no Crime Cibernético

Em 2024, 87% das organizações sofreram algum tipo de ataque envolvendo IA. Esse número alarmante reflete a escala que a IA proporciona aos criminosos, permitindo ataques mais sofisticados, rápidos e personalizados, tornando a cibersegurança um desafio ainda maior.



Capítulo 2: O que é Inteligência Artificial?

Definindo a “Inteligência Artificial”

De forma simplificada, a Inteligência Artificial é qualquer processo computacional que tente simular a inteligência humana. Essa definição abrange desde sistemas básicos até os mais avançados, como os modelos generativos.

Modelos de IA

Existem dois principais tipos de Inteligência Artificial:

- **IA Tradicional:** Baseada em regras e algoritmos simples, como o corretor automático de texto do seu celular, que aprende com o comportamento do usuário.

- **IA Generativa:** Capaz de gerar conteúdos novos a partir de uma base de dados, como textos, imagens ou vídeos. O ChatGPT, por exemplo, utiliza uma vasta base de dados de textos para gerar respostas com qualidade humana. Porém, vale ressaltar que esse modelo ainda não é “terminator”. Ou seja, ele não possui auto-consciência e não é capaz de evoluir sozinho.

Neste eBook, focaremos na IA generativa e seu impacto na segurança cibernética.

Capítulo 3: Tipos de IA Generativa Mais Comuns

Modelos de Linguagem em Grande Escala (LLMs)

Os LLMs, como o ChatGPT e o Copilot, **geram textos com características semelhantes a escrita humana** e são capazes de interpretar, resumir e comparar contextos. No entanto, criminosos os utilizam para criar **phishing personalizado**, explorando informações específicas da vítima.

Por exemplo, um atacante pode buscar informações mais precisas sobre a vítima nas redes sociais e utilizar uma IA de modelo LLM para gerar um texto convincente e humanizado, como: "Parabéns, Thiago, por participar da corrida X! Clique aqui para ver sua foto". Esse tipo de mensagem aumenta significativamente a chance de a vítima clicar no link malicioso.

Modelos de Difusão

Os modelos de difusão **geram imagens realistas a partir de descrições**. Um exemplo é a criação de documentos falsos, como uma carteira de motorista da Flórida, que pode ser usada para burlar sistemas de verificação em bancos digitais.

Rede Adversarial Generativa (GANs)

As GANs consistem em duas IAs competindo entre si: uma produz e a outra verifica sua qualidade. Dessa forma, elas são **capazes de criar conteúdos de áudio e vídeo perfeitos**. Consequentemente, os criminosos as utilizam para criar deepfakes, como vídeos falsos de executivos solicitando transferências financeiras, enganando até mesmo equipes financeiras experientes.

Capítulo 4: Como a IA se Torna uma Arma Cibernética

Reconhecimento e Automação

A IA permite analisar grandes volumes de dados para **identificar alvos e vulnerabilidades com eficiência inigualável,** automatizando o processo com eficiência muito superior à humana.

Melhora Personalização e Alvos

Quando um ataque não é bem-sucedido, a IA pode se adaptar, ajustando sua abordagem para superar as barreiras que ela identificou. Por exemplo, se um e-mail de phishing não funcionar, ela pode criar outro com um texto mais apelativo, superando barreiras de defesa.

Automatização e Adaptação de Ataques

Quando um ataque não é bem-sucedido, a IA pode se adaptar, ajustando sua abordagem para superar as barreiras que ela identificou. Por exemplo, se um e-mail de phishing não funcionar, ela pode criar outro com um texto mais apelativo, superando barreiras de defesa.

Capítulo 5: Outras Formas de Uso Malicioso da IA

Data Poisoning - Envenenamento de Dados

Como mencionamos anteriormente, a IA Generativa consulta um banco de dados para formular as suas respostas. **O Data Poisoning ocorre quando esses dados de consulta são contaminados.** Por exemplo, a Microsoft havia criado um chatbot para interagir com os usuários. Porém, criminosos conseguiram imputar dados e treinar os robôs para que tivessem resultados racistas e ofensivos.

Engenharia Social e Vazamento de Senhas

A engenharia social **é um dos principais problemas enfrentados atualmente** pois explora vulnerabilidades humanas para obter acesso a um sistema protegido ou de informações confidenciais. Com acesso a dados pessoais, como time de futebol, nome dos filhos ou datas de nascimento, os criminosos conseguem adivinhar senhas com mais facilidade. **Estudos mostram que, com contexto, 16% das senhas podem ser quebradas em apenas 12 horas.**

Desenvolvimento de Malware

A Inteligência Artificial também é utilizada por atacantes para desenvolver malwares mais sofisticados, como o MGT, que gera milhares de malwares com assinaturas diferentes, dificultando sua detecção por sistemas de proteção.

Capítulo 6: Casos Reais de Uso Malicioso da Inteligência Artificial

- **Deepfake de Brad Pitt:** Golpistas usaram vídeos falsos de Brad Pitt para convencer uma vítima a transferir 800 mil euros (cerca de 5 milhões de reais), alegando que o ator precisava de ajuda financeira.

Golpista usa imagens falsas de Brad Pitt para enganar francesa e roubar US$ 1,2 milhão

A reação online ridicularizando a credulidade da mulher resultou na decisão da TF1 de retirar o programa

Por equipe do National Post
Publicado em 14 de janeiro de 2025 | Última atualização em 22 de março de 2025 | Leitura de 3 minutos | Junte-se à conversa

Uma das várias imagens deepfake de Brad Pitt usadas para enganar uma francesa em US$ 1,2 milhão. FOTO: TF1

Um falso Brad Pitt gerado por computador e usado por um golpista enganou um designer de interiores francês de 53 anos, fazendo-o pensar que estava hospitalizado com câncer renal e precisava de fundos para pagar o tratamento.

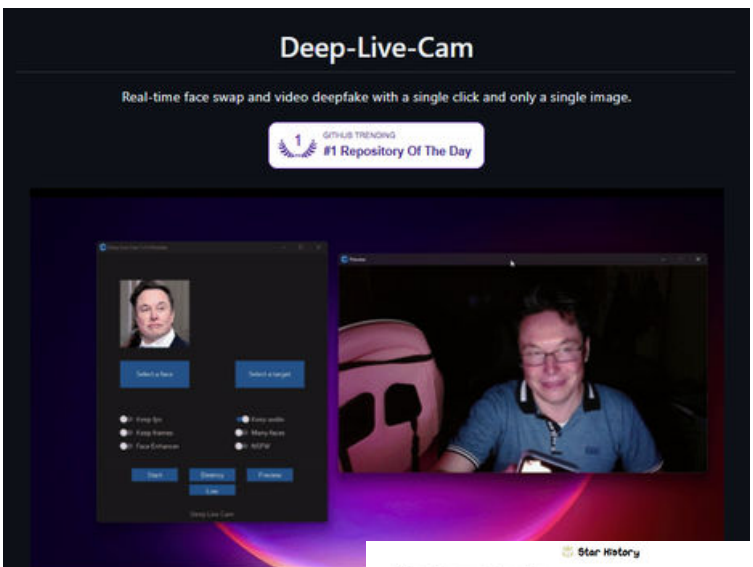
Imagem gerada por golpista para simular uma foto de Brad Pitt hospitalizado

Capítulo 6: Casos Reais de Uso Malicioso da Inteligência Artificial

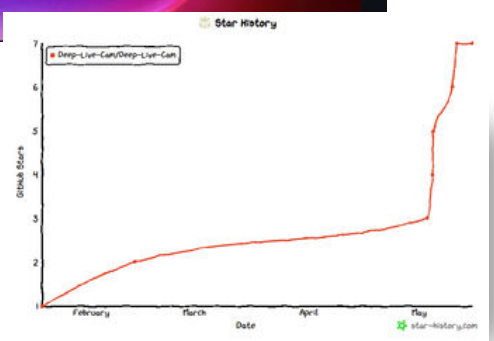
- **Imagens diversas e vídeos gerados por IA avançada**

Capítulo 6: Casos Reais de Uso Malicioso da Inteligência Artificial

O repositório Deep-Live-Cam se tornou um dos mais populares do GitHub, destacando o crescente interesse de hackers em ferramentas de deepfake.



https://github.com/hacksider/Deep-Live-Cam



https://www.star-history.com/#Deep-Live-Cam/Deep-Live-Cam&Date

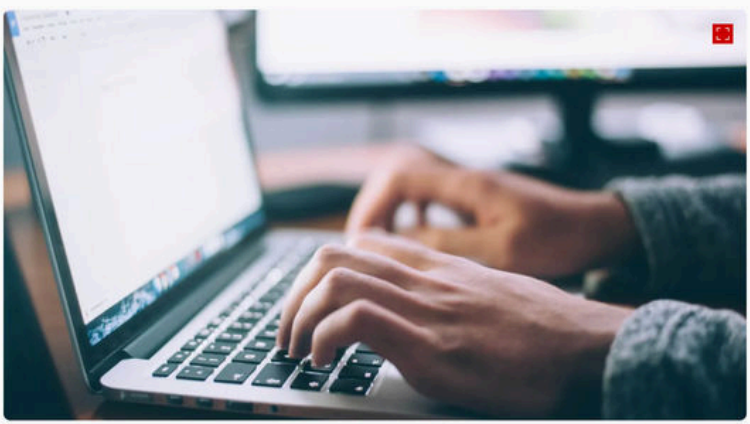
Capítulo 6: Casos Reais de Uso Malicioso da Inteligência Artificial

- **Fraude em Hong Kong:** Criminosos usaram deepfakes em uma videoconferência para personificar executivos de uma empresa, convencendo um funcionário a **transferir 25 milhões de dólares para uma conta fraudulenta.**

Funcionário de multinacional paga US$ 25 mi a golpista que usou "deepfake" para simular reunião

As autoridades estão cada vez mais preocupadas com o potencial prejudicial representado pela tecnologia de inteligência artificial

Heather Chen e Kathleen Magramo, da CNN
04/02/2024 às 10:02



Na videoconferência descobriu-se que todos os participantes eram 'fakes' • Unsplash

ouvir notícia

Um funcionário financeiro de uma empresa multinacional pagou US$ 25 milhões a fraudadores que usaram tecnologia deepfake para se passar por diretor financeiro da empresa em uma videoconferência, segundo a polícia de Hong Kong.

O golpe fez com que o trabalhador fosse levado a participar de uma vídeo chamada com o que ele pensava que fossem vários outros membros da equipe, mas todos na verdade eram criações falsas, disse a polícia de Hong Kong em uma coletiva de imprensa na última sexta-feira (02).

Capítulo 7: O Futuro da IA no Crime Cibernético

Ameaças Internas Avançadas

Deepfakes realistas já estão sendo usados para conseguir empregos falsos, permitindo que criminosos acessem redes corporativas. No futuro, a IA poderá realizar esse processo de forma autônoma.

Agentes Autônomos

Ferramentas como o ChatGPT, em versões mais avançadas, podem realizar tarefas complexas de forma autônoma, como planejar viagens. Criminosos podem usar essa capacidade para automatizar ataques no futuro.

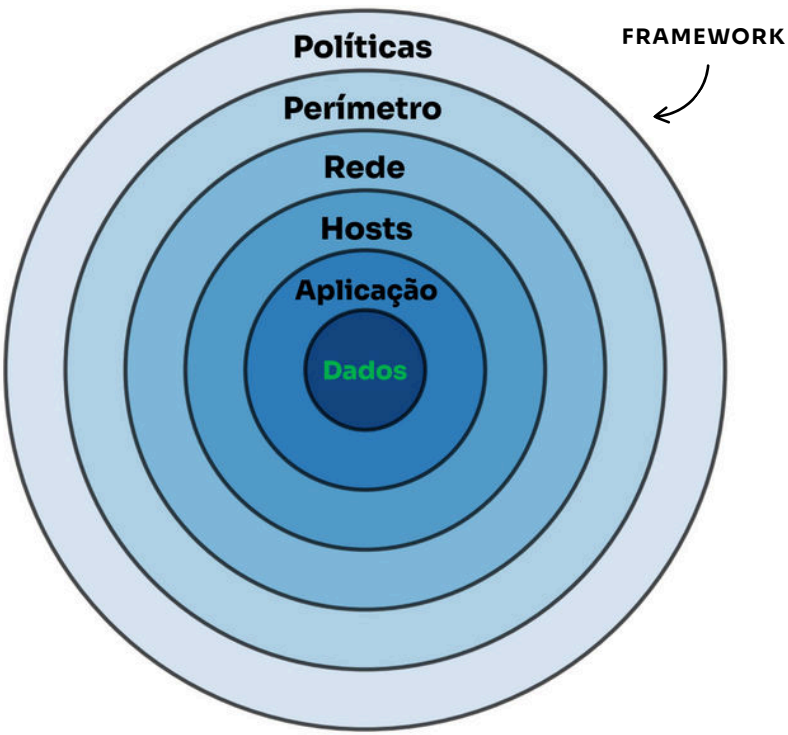
Desenvolvimento Autônomo de Malware

Em breve, a IA poderá buscar vulnerabilidades, desenvolver códigos de ataque e executá-los de forma totalmente autônoma, aumentando a velocidade e a eficiência dos crimes cibernéticos.

Capítulo 8: Como se Proteger?

Camadas Fundamentais de Defesa

A defesa contra ataques potencializados por IA começa com o básico bem feito. As camadas essenciais de um framework de segurança incluem:



- **Políticas e Governança:** Utilize a Inteligência Artificial para auxiliar na definição de políticas claras de segurança, como gerenciamento de senhas e resposta a incidentes.

- **Perímetro de Rede:** Utilize firewalls e VPNs, potencializados por IA para conseguir identificar ameaças em tempo real e ir se adaptando e aprimorando a segurança.

- **Rede Interna:** Implemente sistemas de detecção de intrusos (IDS) e filtros de e-mail, com IA para detectar anomalias e comportamentos que escapam de sistemas tradicionais.

- **Hosts:** Use soluções de EDR (Endpoint Detection and Response) e gerenciamento de patches para corrigir vulnerabilidades.

- **Aplicações:** Adote autenticação multifator (MFA) e revisão de código, com IA para identificar riscos e adaptar os sistemas de autenticação em tempo real.

- **Dados:** Implemente gestão de identidade, DLP (Data Loss Prevention) e controles de acesso adaptativos, com IA para classificar dados automaticamente.

VOCÊ SABIA?

A IA também pode ser uma aliada na defesa, ajudando a:

- Criar políticas personalizadas com base no risco da empresa.
- Detectar anomalias em redes e endpoints.
- Identificar comportamentos maliciosos antes de ataques.
- Adaptar sistemas de autenticação em tempo real.

Capítulo 9: Recuperação e Resiliência

A Realidade dos Ataques Cibernéticos

Hoje, ser alvo de um ataque cibernético não é mais uma questão de "se", mas de "quando". **A IA aumenta a velocidade e o alcance dos ataques**, tornando a defesa cada vez mais desafiadora. Por isso, **a recuperação pós-ataque é tão importante quanto a prevenção.**

Segundo o **Relatório Anual da Veeam sobre Ransomware Trends** constata que o tempo médio para a restauração de todos os ambientes tecnológicos de uma empresa é de 21 dias.

Os cibercriminosos compreendem que empresas que possuem um backup bem estruturado, não precisarão pagar pelo resgate das informações. Por essa razão 97% dos ataques ocorridos em 2024 miraram backups, o que reflete ainda mais a necessidade urgente de possuir uma estratégia robusta de proteção, resiliência de dados e disaster recovery.

Capítulo 9: Recuperação e Resiliência

Estratégia de Resiliência de Dados

- **Políticas de Backup:** Adote uma política robusta, como a **"Golden Rule" 3-2-1-1-0 (3 cópias, 2 mídias diferentes, 1 off-site, 1 imutável, 0 erros)** que deve ser testada regularmente para garantir a recuperação dos dados.

Se você deseja se aprofundar na Regra 3-2-1-1-0, acesse o conteúdo completo no nosso blog.

Capítulo 9: Recuperação e Resiliência

- **Plano de Disaster Recovery:** Defina tempos de recuperação (RTO) e pontos de recuperação (RPO) para sistemas críticos, negociando com o negócio.

Mas o que é RPO e RTO ?

O Recovery Point Objective (RPO) define o intervalo máximo aceitável de perda de dados em caso de falha, indicando quanto tempo de informações uma empresa pode perder sem comprometer a continuidade dos negócios. Quanto menor o RPO, maior a frequência dos backups, reduzindo a perda de dados em caso de desastre.

O Recovery Time Objective (RTO), por sua vez, determina o tempo máximo que um sistema pode ficar indisponível antes de causar impactos críticos. Quanto menor o RTO, maior a necessidade de tecnologias rápidas de recuperação, como failover automático e storage redundante, para garantir a rápida retomada das operações.

RPO e RTO na Prática		
Cenário	**RPO (Ponto de Recuperação)**	**RTO (Tempo de Recuperação)**
Banco de Dados Financeiro	5 minutos	15 minutos
E-commerce	10 minutos	30 minutos
Sistema de Atendimento ao Cliente	30 minutos	1 hora
Arquivos Internos e E-mail	12 horas	6 horas

Saiba mais sobre RPO e RTO em nosso Blog. VER ARTIGO

Capítulo 10: Perguntas e Respostas

Como É Feita a Inserção de Dados na IA para Criar Malware?

Existem diversas técnicas para isso. Criminosos inserem dados maliciosos na base de uma IA para moldar seu comportamento, como no caso do chatbot Tay da Microsoft, que foi treinado a exibir comportamentos ofensivos.

Há Riscos de Compartilhamento de Dados com IA?

Cada ferramenta de IA tem termos de uso específicos sobre o uso de dados. Recomendamos revisar esses termos, mas a maioria das IAs utiliza dados de forma anonimizada para aprendizado ou estatísticas. Evite inserir informações confidenciais em IAs públicas.

A IA Pode Ajudar na Proteção?

Sim, a IA pode ser usada em ferramentas de segurança para detectar anomalias, analisar comportamentos e adaptar defesas em tempo real, como no Veeam, que utiliza IA para proteger backups.

Capítulo 11: A Solução da Penso

COMO A PENSO PODE AJUDAR A SUA EMPRESA?

- ✓ Realizar BIAs para identificar processos críticos.
- ✓ Construir políticas globais e contínuas de resiliência de dados.
- ✓ Implementar disaster recovery com **recuperação de 15 minutos.**
- ✓ Garantir governança e testes periódicos.
- ✓ Fazer um diagnóstico da sua infraestrutura atual

O que é o BIA? Business Impact Analysis

Realizamos uma análise de impacto de negócio (BIA) para identificar sistemas e processos críticos, definindo tempos aceitáveis de interrupção do seu negócio.

Implementar Estratégia de Resiliência de Dados Global e Contínua

Auxiliamos os nossos clientes a construírem uma política sólida de backup global e contínua, envolvendo todos os dados da empresa, mesmo em ambientes multicloud.

Resiliência de Processos e Sistemas (DR)

Implementamos soluções de Disaster Recovery, estabelecendo em conjunto com as demais áreas de negócio quais serão as prioridades e quanto tempo levará para recuperação, além de realizarmos testes periódicos para garantir o funcionamento perfeito.

PRONTO PARA FORTALECER A RESILIÊNCIA DOS SEUS DADOS?

Transforme a tecnologia em aliada da sua empresa

veeam
Competency Partner
DRaaS
BaaS *for 365*
Off-site Backup

Penso Tecnologia recebe prêmio da Veeam como "Impact VCSP Partner of the Year"

Reconhecimento destaca a excelência da Penso Tecnologia em resiliência de dados e consolida sua posição como referência no mercado de TI

Por PressWorks
14/03/2025 08h40 · Atualizado há 4 semanas

Conheça nosso portfólio completo para proteger dados, otimizar processos e garantir a continuidade do seu negócio.

PROTEÇÃO DE DADOS

Veeam Backup

Solução para proteger os dados da sua empresa.

Backup na Nuvem

Segurança para sua empresa e benefícios aos usuários.

Disaster Recovery

Proteção contínua para uma empresa mais segura.

Backup 365

Backup seguro dos seus e-mails Microsoft 365.

CYBER SECURITY

Pentest como serviço

Mapeamento e redução de falhas na segurança.

Segurança para o usuário

Proteção efetiva contra ameaças virtuais.

Firewall como serviço

Proteção avançada da rede da sua empresa.

SERVIÇOS DE TI

Gestão e suporte de TI

Atendimento especializado remoto ou presencial.

Suporte Avançado N2, N3 e Especialistas

Apoio para demandas de alta complexidade

COMPUTAÇÃO EM NUVEM

Penso Cloud Corporativo

Migração segura para nossa nuvem

Penso S3 Storage

Armazenamento em nuvem

EMAIL E COLABORAÇÃO

PensoMail

E-mail corporativo personalizado para sua empresa.

Zimbra: e-mail corporativo

Conheça nossas soluções baseadas em Zimbra

Auditoria de e-mail

Acompanhe os e-mails trafegados no seu negócio.

Cloud Antispam

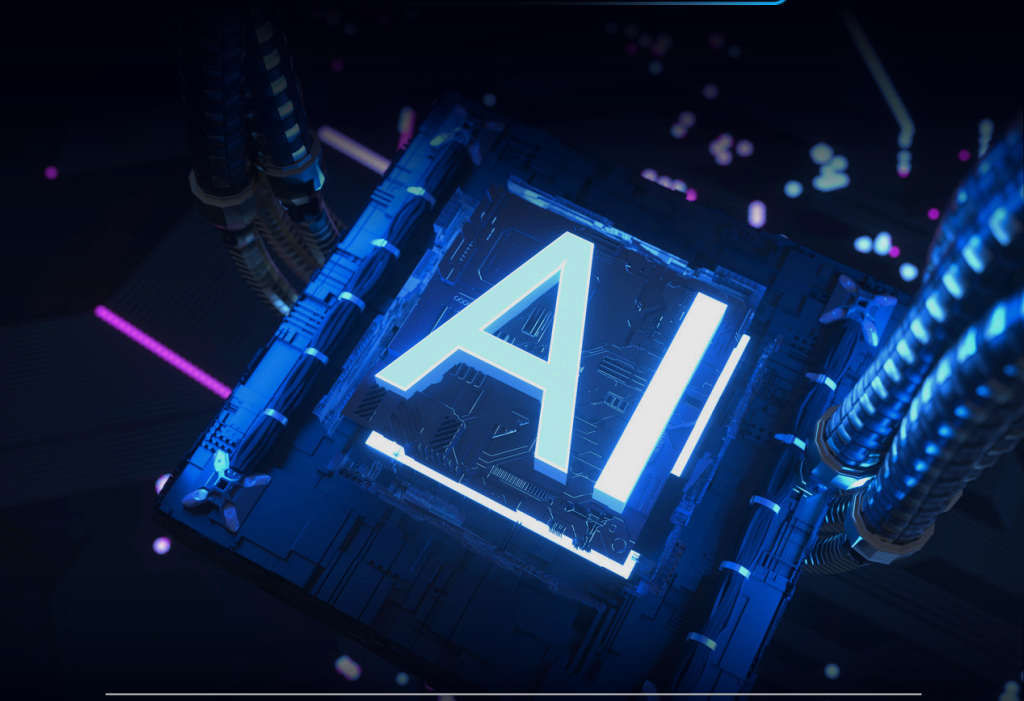
A solução antispam ideal para nuvem.

Conclusão

A IA generativa é uma ferramenta poderosa, mas também uma arma perigosa nas mãos de criminosos. Deepfakes, phishing personalizado e engenharia social assistida por IA são apenas algumas das ameaças que enfrentamos.

No entanto, com o básico bem feito, políticas robustas de backup e disaster recovery, e o apoio de parceiros como a Penso e a Veeam, é possível se proteger e se recuperar de ataques.

Esteja preparado: em um cenário de guerra cibernética, a resiliência é a chave para a sobrevivência.



Penso

Quando a IA se torna uma arma arma cibernética

#Obrigado

Apoio: veeam Adistec