Riscos cibernéticos e inteligência artificial

05.10.2023 4 minutos de leitura

Autores

Felipe Palhares

Áreas relacionadas

Proteção de Dados e Cybersecurity

Assuntos

Inteligência Artificial Proteção de Dados e Cybersecurity

Imagine o seguinte cenário: você, diretor financeiro de uma empresa, recebe uma ligação da presidente da organização. A chamada aparece no seu dispositivo como sendo da linha pessoal da presidente. A voz do outro lado da ligação é igual à da presidente, inclusive o timbre, entonação, as palavras utilizadas e a forma de falar.

O objetivo da ligação é solicitar que uma transferência financeira seja feita a um novo fornecedor, para uma demanda urgente, que precisa ser iniciada o quanto antes. Como a maioria das pessoas nessa situação, você faz a transferência. Mas do outro lado da linha não era a presidente que falava, e sim uma voz manipulada para parecer exatamente como a dela.

Essa situação não é meramente hipotética. Tem acontecido de forma cada vez mais frequente, nas mais diversas empresas, em cenários muito similares ao apresentado acima. Por meio de técnicas de mascaramento de linhas telefônicas ou de e-mails, e pela utilização de ferramentas de inteligência artificial, golpes como esse se tornaram corriqueiros, e cada vez mais elaborados.

Como não seria possível esgotar todos os riscos cibernéticos trazidos pela inteligência artificial neste artigo, abaixo serão apresentados 3 riscos relevantes, que devem estar no radar de toda e qualquer empresa nos dias de hoje.

1. *Deepfakes*

Sistemas de inteligência artificial aprendem com dados e são excelentes em replicar dados que lhes são apresentados. Por essa mesma razão, são ferramentas excepcionais para criar réplicas de imagens, vídeos ou de voz. Dentre essas atividades está a criação de *deepfakes*, representações ultrarrealistas de uma determinada pessoa, como o cenário da presidente da empresa acima.

Criminosos vêm se utilizando de *deepfakes* para enganar pessoas ou para extorquir vítimas, manipulando fotos e vídeos de modo a fazer com que terceiros tenham a plena convicção de que um determinado indivíduo está realizando um ato que nunca realizou, como uma solicitação para que uma transferência financeira seja feita.

Com o advento dos *deepfakes*, torna-se cada vez mais relevante que empresas criem políticas e estratégias internas para autenticar os seus colaboradores, especialmente aqueles que ocupem posições elevadas dentro da organização.

2. Compartilhamento de Informações

Na era do ChatGPT e de outras ferramentas similares, que fornecem informações de acordo com dados que são inseridos pelos usuários (inclusive trechos de documentos), é fundamental ter no radar que informações confidenciais da empresa podem estar sendo compartilhadas pelos empregados sem que estes tenham ciência dos riscos do uso dessas ferramentas.

Considerando que os grandes modelos de linguagem são treinados com um volume enorme de dados, é natural que eles guardem todas as informações que sejam apresentadas pelos usuários para utilizá-las posteriormente quando forem úteis para a resposta de questionamentos apresentados por outros usuários.

Em outras palavras, os trechos de documentos que estão sendo inseridos por colaboradores em sistemas de inteligência artificial, no intuito de obtenção de respostas melhores às suas solicitações (como, por exemplo, um pedido de revisão da gramática de um documento), podem representar um incidente de segurança, quebrando a confidencialidade daquele documento.

3. *Plugins* de Terceiros

Diversas soluções de inteligência artificial permitem que o usuário utilize *plugins* desenvolvidos por terceiros para se conectar às soluções, ou integrá-las de forma distinta da inicialmente concebida pelos provedores da inteligência artificial.

Terceiros mal-intencionados podem explorar vulnerabilidades desses plug-ins para ter acesso aos comandos realizados pelos usuários, incluindo toda e qualquer informação fornecida pelo usuário ou gerada como resposta pela inteligência artificial. Em alguns casos, o terceiro pode conseguir interceptar a resposta, modificando os seus termos, de modo a causar, além da quebra de confidencialidade, uma quebra à integridade da informação.

Além disso, diante do crescimento exponencial do interesse por soluções de inteligência artificial, criminosos podem desenvolver os seus próprios plug-ins e divulgá-los como se fossem códigos benignos, sem que os usuários inicialmente saibam do seu intuito malicioso.

Esses são somente alguns riscos cibernéticos advindos do uso de sistemas de inteligência artificial. Nesse cenário, é primordial que as empresas estejam preparadas para endereçar tais riscos, evitando que incidentes de segurança ocorram por desconhecimento ou falta de atuação de medidas prévias.

› PROFISSIONAIS

Profissionais que aliam visão jurídica e visão de mercado

Felipe Palhares