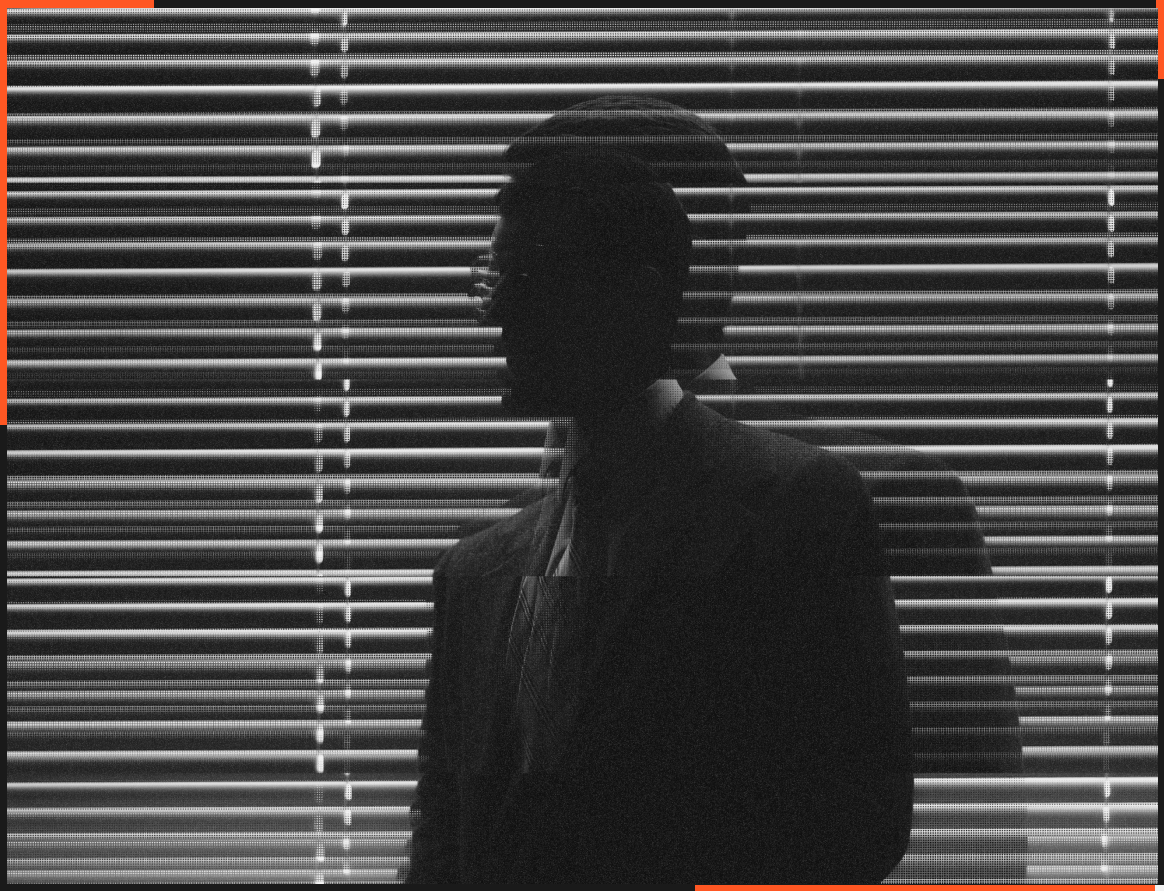


E-book

Deepfakes em campanhas de fraude corporativa

Como a simulação de executivos se tornou uma ameaça direta à reputação e segurança das empresas

AXUR

Sumário Executivo

O avanço das tecnologias de inteligência artificial generativa — especialmente os deepfakes — trouxe à tona uma ameaça invisível, porém poderosa: a manipulação da identidade de líderes e figuras de autoridade como arma de fraude.

Executivos de alto escalão se tornaram alvos estratégicos em campanhas que exploram uma nova dimensão de vulnerabilidade: a sua imagem pública.

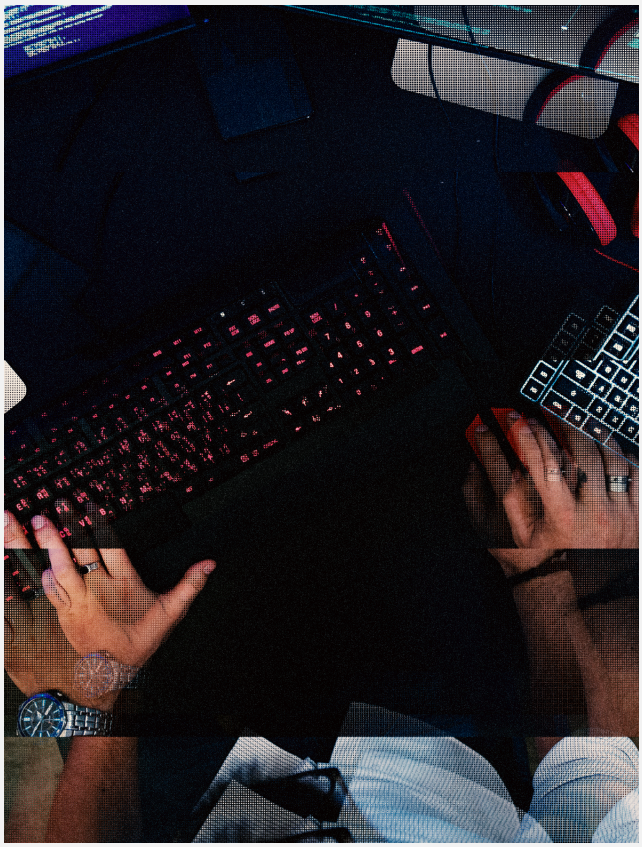
De chamadas de vídeo com voz clonada a perfis falsos em redes sociais, os golpistas estão ultrapassando barreiras tecnológicas e psicológicas para enganar colaboradores, movimentar grandes somas financeiras e comprometer a reputação das organizações.

Neste eBook, exploramos como essas campanhas são estruturadas, por que os executivos estão no centro da mira, e como as empresas podem se proteger. Com base em nosso framework proprietário Fraud Neuron, decodificamos os padrões e táticas por trás dessas operações e mostramos como a inteligência artificial pode ser usada tanto para o ataque quanto para a defesa.

Deepfakes: a ilusão perfeita a serviço da fraude

Os deepfakes deixaram de ser curiosidades tecnológicas ou ameaças restritas a celebridades. Hoje, eles alimentam fraudes corporativas em escala, usando a imagem e a voz de executivos para convencer colaboradores a tomar decisões críticas — como autorizar transferências bancárias, abrir acessos a sistemas ou divulgar informações estratégicas.

Com uma única imagem pública e poucos segundos de áudio, criminosos conseguem gerar avatares realistas que falam, gesticulam e expressam emoções de forma convincente. A natureza dessas campanhas é estratégica: elas se apoiam em engenharia social, na autoridade da figura executiva e em situações de urgência para comprometer a racionalidade da vítima.



Por que executivos são os novos alvos preferenciais

Não é coincidência que CFOs, diretores regionais e CEOs estejam no centro dessas fraudes. Eles reúnem três características que os tornam valiosos para operadores maliciosos:

1. Alto nível de exposição pública: entrevistas, vídeos institucionais, eventos online e redes sociais fornecem material farto para treinamento de modelos de deepfake.
2. Autoridade inquestionável: um comando vindo do “CFO” por vídeo ou áudio não costuma ser questionado por colaboradores, especialmente sob pressão.
3. Acesso a recursos críticos: os executivos geralmente detêm ou influenciam decisões financeiras e operacionais sensíveis.

Com essas variáveis combinadas, campanhas bem orquestradas conseguem contornar controles técnicos, explorando uma vulnerabilidade essencialmente humana: o respeito à hierarquia. Além de fraudes financeiras diretas, há indícios de uso dessas técnicas para ganho de acesso inicial em ambientes corporativos via social engineering, ou engenharia social, dirigida.

Anatomia de um deepfake: além da ilusão digital

Para entender a complexidade e o impacto desse fenômeno, utilizamos aqui a estrutura tática de classificação baseada no Fraud Neuron, framework proprietário da Axur desenvolvido para descrever, categorizar e monitorar fraudes digitais complexas. A seguir, descrevemos cada uma das fases típicas de uma campanha com deepfakes, conforme observado em incidentes reais.

→ Identificação de alvos (T1000) – Target Identification

A campanha maliciosa realiza uma identificação precisa de alvos, priorizando tanto indivíduos quanto organizações de alto valor. Entre os **alvos individuais (TQ1110)**, destacam-se **funcionários (P1112)** com acesso a finanças ou operações críticas e **executivos seniores (P1113)**, como CFOs, que são utilizados como modelo para os deepfakes. No espectro corporativo (TQ1120), os atores visam **grandes corporações (P1122)** com estruturas descentralizadas e cadeias de comando fragmentadas, o que facilita a engenharia social. Também há indícios de exploração de **instituições financeiras (P1123)**, especialmente no uso da identidade de executivos de empresas do setor cripto.

→ Temas (T2000) – Themes

A narrativa das fraudes é cuidadosamente construída com temas financeiros e de oportunidades profissionais. Atores utilizam ofertas de investimento (P2113) como pretexto, por exemplo, através de deepfakes de executivos da Binance promovendo supostas oportunidades cripto. Em paralelo, temas de emprego (TQ2200) são explorados por meio de falsas ofertas de trabalho (P2211) veiculadas em redes sociais, com vídeos deepfake de celebridades brasileiras promovendo produtos ou serviços inexistentes. Esses temas têm alto apelo e ampliam o alcance das campanhas fraudulentas.

→ Reconhecimento (T3000) – Reconnaissance

A fase de reconhecimento envolve extensa coleta de dados públicos (TQ3110). Os atacantes extraem informações visuais e vocais de fontes abertas (P3111) como entrevistas e apresentações públicas, além de colher dados de redes sociais (P3113) para entender linguagem corporal, padrões de fala e contexto organizacional. Esses dados alimentam modelos de AI capazes de gerar deepfakes convincentes. Não há indícios de uso de fontes privadas ou técnicas de coleta técnica, mas a profundidade do conteúdo gerado sugere uma abordagem meticulosa de *open-source intelligence*.

→ Recursos (T4000) – Resources

Os operadores implementam infraestrutura dedicada para suas fraudes. Eles criam contas falsas em redes sociais (P4113) para simular a presença de figuras públicas ou executivos. Domínios são registrados (P4211) com nomes similares às marcas corporativas para hospedar conteúdo deepfake ou páginas de phishing. A campanha utiliza infraestrutura de servidores (P4212) para armazenar vídeos manipulados e realizar transmissões em tempo real. Carteiras de criptomoedas (P4214) são preparadas para receber fundos das fraudes. A operação é viabilizada por sistemas de inteligência artificial (P4311) para geração dos vídeos e softwares especializados (P4312) que realizam lip-sync, clonagem de voz e manipulação de áudio com alta fidelidade.

→ Conversão (T7000) – Conversion

A extração de valor ocorre por meio de transferências bancárias fraudulentas (P7111) após convencimento da vítima. Em outros casos, o valor desviado é convertido e movimentado via transações em criptomoedas (P7113), dificultando o rastreamento. Essa tática também se aplica a golpes com falsas promoções e vendas, onde o pagamento é feito diretamente para carteiras digitais controladas pelos operadores da campanha.

→ Simulação de identidade (T5000) – Identity Simulation

A simulação de identidade é o núcleo da campanha. Os atacantes praticam personificação de funcionários (P5111), sobretudo executivos de alto escalão, em chamadas de vídeo e áudios. Também utilizam deepfakes de personalidades públicas (P5113) para gerar vídeos promocionais falsos que circulam como anúncios pagos. Para isso, aplicam técnicas de spoofing de canal (P5213), simulando chamadas por Zoom, Microsoft Teams e outras plataformas legítimas, criando uma ilusão de autenticidade com uso sincronizado de imagem e voz. A junção destas técnicas resulta em engajamento direto com as vítimas em tempo real ou por meio de vídeos convincentes.

→ Engenharia Social (T6000) – Social Engineering

A campanha é fundamentada em engenharia social sofisticada. Os operadores criam situações de emergência (P6112) onde o deepfake de um executivo exige ação imediata, geralmente uma transferência bancária ou liberação de acesso. Empregam pressão por tempo (P6121) para reduzir a capacidade crítica das vítimas, forçando decisões rápidas. Ao utilizar o canal corporativo e a imagem de autoridade, os atacantes induzem pressão social (P6123), fazendo com que a vítima se sinta obrigada a cumprir as ordens do "superior". A manipulação psicológica é aumentada pelo realismo da imagem e da voz geradas por IA.

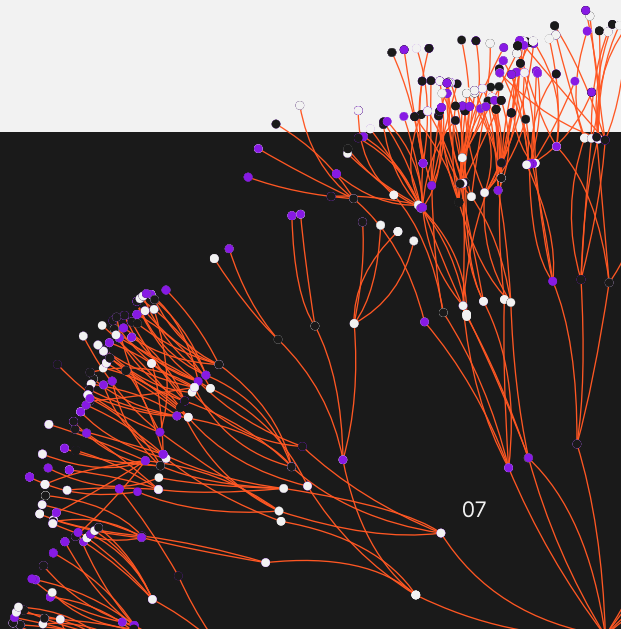
→ Impactos (T8000) – Impacts

Os impactos incluem perdas financeiras diretas (P8111), como no caso documentado de US$25 milhões desviados por meio de uma videoconferência com deepfake do CFO. Há também danos à reputação (P8211) quando empresas têm suas marcas exploradas em campanhas fraudulentas, como nos vídeos falsos veiculados com celebridades brasileiras. Por fim, a exposição a esses enganos causa impacto psicológico (P8213) nos colaboradores enganados, que acreditaram estar interagindo com seus superiores ou com figuras públicas reais.

Fraud Neuron

Saiba mais sobre o Fraud Neuron e contribua para o mapeamento das fraudes digitais:

github.com/axur/fraudneuron



DeepFake as a Service

O modelo Deepfake-as-a-Service consiste na comercialização direta de tecnologias de mídia sintética baseadas em inteligência artificial, especialmente focadas na criação de vídeos, áudios e avatares deepfake sob demanda. Trata-se de uma estrutura de negócio típica, onde fornecedores oferecem acesso simplificado e de baixo custo a tecnologias complexas para clientes sem expertise técnica avançada.

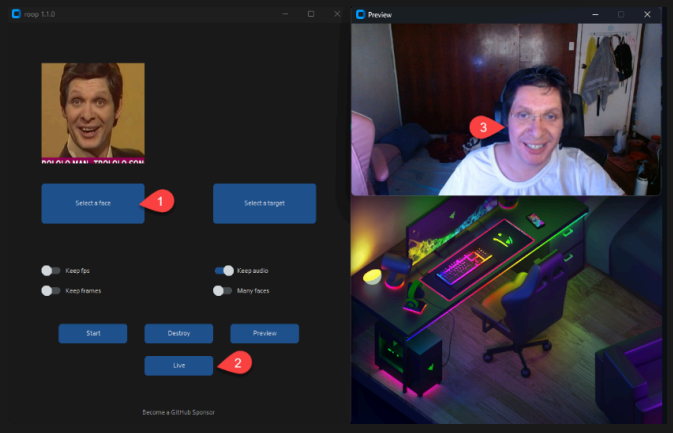
A característica essencial desse modelo é a disponibilização de plataformas online, nas quais compradores solicitam serviços personalizados, tais como criação de vídeos manipulados, chamadas em tempo real com avatares deepfake, ou replicação de voz para campanhas específicas de fraude.

O mercado de DeepFake as a Service (DFaaS) cresceu rapidamente nos últimos anos, impulsionado pela acessibilidade e simplicidade das ferramentas baseadas em inteligência artificial generativa. Neste modelo de negócio, plataformas comerciais oferecem tecnologias avançadas que permitem a criação de conteúdos manipulados realistas com facilidade, mesmo para usuários sem experiência técnica significativa.

Ao reduzir drasticamente as barreiras técnicas e financeiras, essas ferramentas não apenas democratizaram o acesso a tecnologias sofisticadas, mas também ampliaram significativamente o risco de uso malicioso das mídias sintéticas.

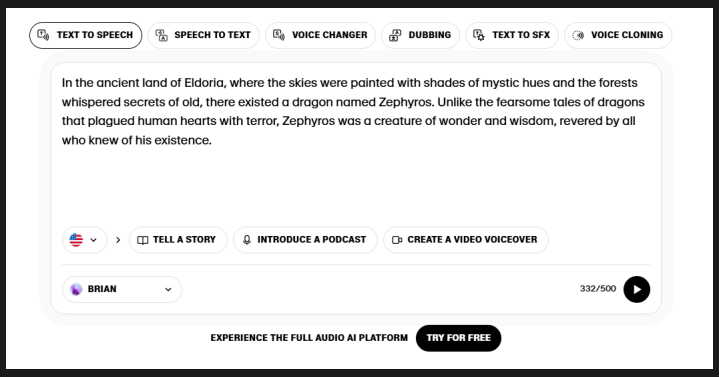
Uma ferramenta notável nesse segmento é o Roop, uma aplicação de código aberto que se tornou popular por permitir substituir a face de qualquer pessoa em um vídeo com apenas uma única imagem da face desejada, dispensando completamente o treinamento de modelos ou a coleta de grandes conjuntos de dados.

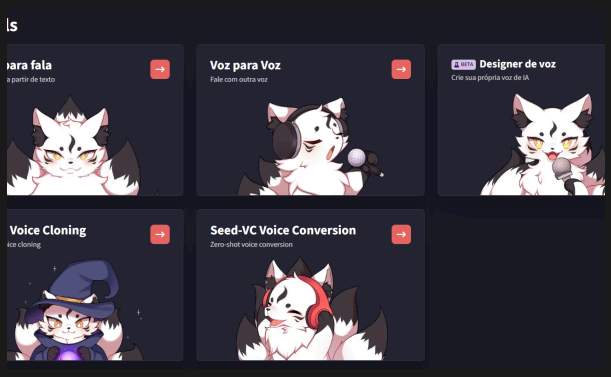
O Roop utiliza redes neurais pré-treinadas para realizar substituições faciais quase instantâneas, produzindo resultados visuais convincentes mesmo com baixo investimento técnico (Roop, GitHub, 2023).



Além disso, existem plataformas que oferecem clonagem de voz e síntese vocal avançada, como a ElevenLabs, que podem ser exploradas de forma maliciosa. Com essa tecnologia, os usuários podem criar áudios realistas a partir de pequenos trechos vocais originais, replicando nuances de emoção, ritmo e entonação com extrema fidelidade.

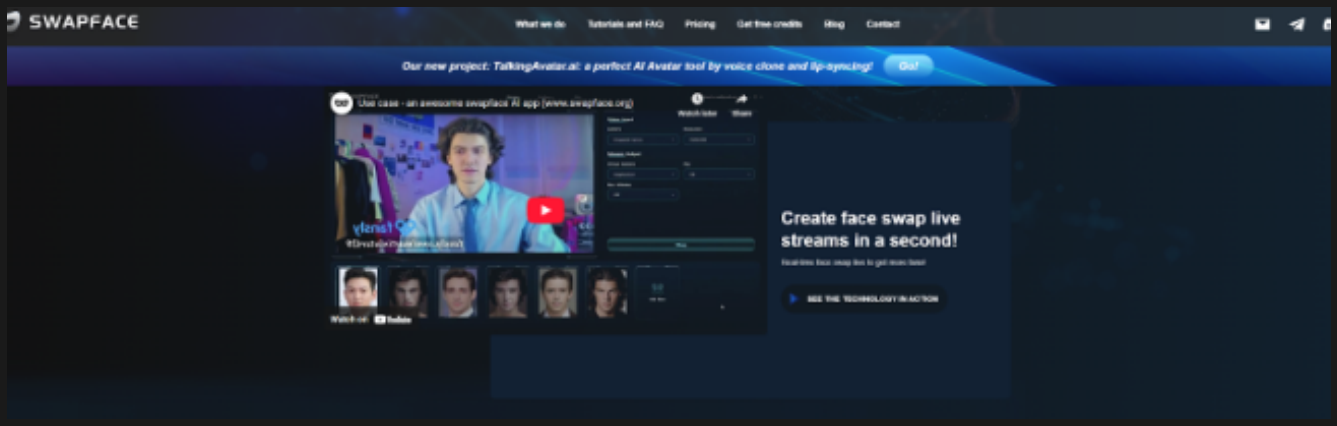
A popularização dessa ferramenta tem sido mencionada como fator decisivo em ataques bem-sucedidos de engenharia social, como os casos recentemente documentados contra executivos de grandes corporações, nos quais a voz clonada foi utilizada para fraudes financeiras significativas (CNN, 2024; Fortune, 2024).





Uma ferramenta semelhante, conhecida como FakeYou, fornece síntese vocal simplificada a partir de texto digitado. FakeYou diferencia-se por possuir um extenso catálogo de vozes sintéticas pré-treinadas, incluindo personagens fictícios e celebridades, facilitando ainda mais o processo de criação de conteúdo sintético.

A simplicidade dessa plataforma a tornou popular não apenas para usos legítimos, mas também como ferramenta rápida para campanhas fraudulentas ou desinformação (FakeYou, 2023).



A ferramenta conhecida como Nudefusion utiliza inteligência artificial para transformar fotos comuns em imagens falsas com conteúdo explícito, em um processo conhecido como "deepnude". Este serviço preocupa especialmente pelo potencial de abuso, invasão de privacidade e difamação. A ferramenta é divulgada de maneira aberta, prometendo gerar resultados rápidos a partir do simples upload de uma imagem comum, dispensando treinamento prévio ou habilidades técnicas específicas por parte do usuário.

Devido ao seu alto potencial de utilização indevida, plataformas similares já foram objeto de investigações e ações legais em diversos países, demonstrando claramente o impacto social e ético das tecnologias de deepfake quando aplicadas para geração não consensual de imagens íntimas falsas.

Ataques que chegaram à mídia: casos reais e técnicas utilizadas

Cenário #1 Sabotagem corporativa

Neste cenário, os criminosos utilizam deepfakes para disseminar desinformação, prejudicar a imagem de marcas, manipular a percepção do mercado ou comprometer negociações de fusões e aquisições (M&A). A sabotagem ocorre através da criação de vídeos e áudios falsos, mostrando executivos ou funcionários em situações comprometedoras ou divulgando informações incorretas sobre produtos e estratégias corporativas.

Caso Real

Um exemplo recente é o ataque contra a empresa Binance, em que criminosos usaram um holograma deepfake do executivo Patrick Hillman, manipulando informações sobre supostos projetos financeiros inexistentes e prejudicando a credibilidade da exchange (Bitdefender, 2023).

Cenário #2 Engenharia social avançada no ambiente corporativo

Nessa situação, o criminoso utiliza deepfake para executar ataques convincentes de engenharia social direcionados a altos executivos ou funcionários estratégicos. O atacante primeiro realiza pesquisas detalhadas sobre executivos, coletando dados pessoais em redes sociais e materiais audiovisuais para treinar modelos avançados. Após obter dados sobre eventos corporativos recentes, como novas joint ventures ou fusões, o atacante cria deepfakes audiovisuais do CEO ou diretor financeiro, utilizando áudio clonado para realizar chamadas fraudulentas.

Exemplo Técnico e Caso Real

Em fevereiro de 2024, criminosos realizaram um ataque sofisticado em Hong Kong usando o deepfake do CFO de uma grande corporação. Uma transferência fraudulenta de mais de US$25 milhões foi autorizada após a vítima receber uma ligação convincente, onde o criminoso explorou informações pessoais do executivo e detalhes específicos da operação financeira (CNN, 2024).

Cenário #3

Ataques de engenharia social contra instituições financeiras

Aqui, o atacante utiliza deepfakes de áudio para comprometer sistemas de autenticação por voz em bancos e instituições financeiras. Após comprar dados pessoais roubados na dark web (nome, CPF, números de contas bancárias) e coletar amostras audiovisuais das vítimas em redes sociais, o criminoso cria clones vocais para enganar sistemas de autenticação, obtendo acesso às contas financeiras dos alvos.

Exemplo Técnico e Caso Real

Embora não haja registro público detalhado específico deste cenário exato, técnicas semelhantes foram identificadas em ataques contra instituições financeiras usando voice cloning para enganar funcionários em centrais de atendimento, permitindo transferências fraudulentas e acesso indevido às contas, demonstrando viabilidade técnica (Fortune, 2024; BleepingComputer, 2024).

Cenário #4

Responsabilidade corporativa e fraudes por Deepfake

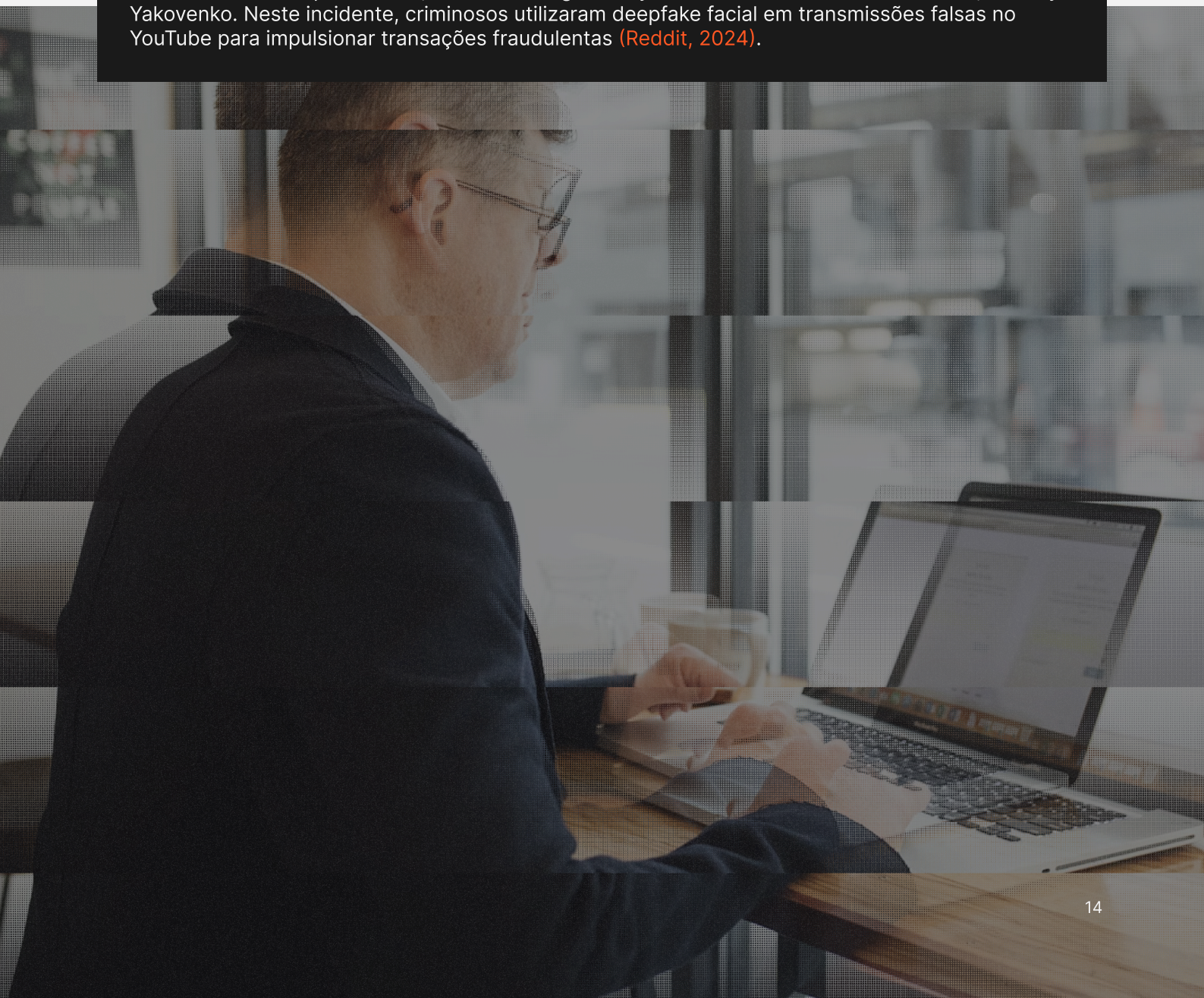
Neste cenário, criminosos criam vídeos deepfakes realistas sugerindo falhas ou acidentes graves envolvendo produtos ou serviços corporativos, visando obter compensação financeira fraudulenta. O criminoso realiza pesquisas detalhadas sobre incidentes anteriores reais e cria uma narrativa visual falsa, utilizando modelos avançados de troca facial para simular vítimas fictícias.

Cenário #5 Manipulação de mercado e ações por Deepfake

Aqui, atacantes utilizam deepfakes para manipular preços de ações no mercado financeiro. O criminoso adquire previamente uma posição acionária, em seguida cria e distribui conteúdo falso em fóruns populares como Reddit ou Stockaholics. Deepfakes convincentes do CEO ou executivos falando sobre eventos financeiros falsos induzem investidores a comprar ou vender ações rapidamente, resultando em grandes variações no preço do ativo.

Exemplo Técnico e Caso Real

Embora casos de deepfake especificamente voltados à manipulação do mercado de ações não tenham sido amplamente divulgados até o momento, cenários semelhantes já foram identificados em fraudes com criptomoedas, como o falso giveaway envolvendo o fundador da Solana, Anatoly Yakovenko. Neste incidente, criminosos utilizaram deepfake facial em transmissões falsas no YouTube para impulsionar transações fraudulentas (Reddit, 2024).



Como identificar e mitigar deepfakes: um guia técnico para empresas e indivíduos

É possível educar o público para identificar deepfakes?

A educação e conscientização da sociedade são cruciais para combater eficazmente a propagação de mídias sintéticas falsas. No entanto, o interesse e o comprometimento público com esse aprendizado podem ser limitados.

Pesquisas realizadas pela PAI apontam que as pessoas preferem não ser instruídas diretamente ou de maneira condescendente sobre o que é verdadeiro ou falso; preferem aprender por meio de contextos e ferramentas que lhes permitam verificar a autenticidade das mídias de forma independente.

Ainda assim, é possível capacitar indivíduos a identificar deepfakes ensinando-os a procurar sinais técnicos claros, tais como:

Sinais que indicam que um vídeo ou imagem pode ser deepfake:

→ Áreas borradas ou desfocadas no rosto, diferentes do restante da imagem.

→ Mudanças repentinas de tom de pele ou iluminação ao redor da face.

→ Bordas duplicadas (ex: duplo queixo, sobrancelhas duplicadas).

→ Falta ou excesso de piscadas; movimentos faciais artificiais.

→ Problemas de sincronização labial e movimentos não naturais.

→ Inconsistências entre o fundo e o objeto principal da cena.

Sinais que indicam que um áudio pode ser deepfake:

→ Frases entrecortadas ou fora de ritmo natural.

→ Mudanças bruscas na entonação da voz.

→ Uso incomum de palavras ou frases que o suposto interlocutor normalmente não diria.

→ Contexto fora de uma discussão habitual ou incapacidade de responder perguntas simples relacionadas.

Sinais que indicam que um texto pode ser manipulado ou fraudulento:

→ Erros ortográficos e gramaticais incomuns.

→ Falta de coerência lógica e fluidez.

→ Email ou número desconhecido ou suspeito.

→ Expressões ou vocabulário incomuns para o suposto remetente.

→ Mensagem fora de contexto ou inesperada.

Entretanto, à medida que a tecnologia avança, distinguir mídias manipuladas das reais torna-se uma tarefa progressivamente complexa. Deepfakes evoluem rapidamente, impulsionados por avanços constantes em modelos generativos e técnicas sofisticadas de inteligência artificial (GANs, autoencoders, entre outros).

Assim, depender exclusivamente da percepção humana para detectar esses conteúdos fraudulentos não é mais suficiente.

Para aumentar a eficácia da identificação, é necessário integrar abordagens técnicas robustas de análise forense digital e algoritmos automatizados.

Como podemos melhorar a capacidade de detecção?

Atualmente, ferramentas como Reality Defender, Microsoft Video Authenticator, Deepware AI, e frameworks acadêmicos como FaceForensics++, estão disponíveis para analisar deepfakes com precisão cada vez maior. Esses sistemas aplicam análises detalhadas utilizando algoritmos de machine learning treinados especificamente para identificar características únicas, tais como inconsistências na iluminação, expressões faciais não naturais, ou artefatos digitais deixados durante o processo de manipulação.

Entretanto, nenhuma ferramenta é totalmente infalível, especialmente contra deepfakes altamente sofisticados e tecnicamente avançados. A eficácia da detecção depende diretamente de treinamento contínuo, ajustes constantes dos modelos e exposição permanente das ferramentas a novos exemplos manipulados.

Além disso, essas ferramentas citadas não realizam um monitoramento avançado para descobrir fraudes usando a imagem de executivos nas redes sociais, sites e plataformas de streaming, o que é essencial para ganhar visibilidade sobre as ameaças.

Como essas ameaças se conectam com proteção de marca e fraude?

Campanhas com deepfakes envolvendo executivos e figuras públicas não causam impacto apenas às vítimas diretas. Quando ocorrem, os efeitos se espalham por toda a estrutura da empresa — atingem a reputação institucional, afetam relações com stakeholders e, muitas vezes, se materializam em prejuízo financeiro. O uso indevido da imagem de líderes organizacionais não é apenas uma violação de identidade individual, mas uma porta de entrada para fraudes com alcance corporativo.

Uso de identidade de executivos em promoções falsas ou campanhas enganosas

Nesse tipo de golpe, um executivo conhecido aparece em vídeo ou áudio promovendo investimentos, anunciando parcerias ou oferecendo vantagens — que, na realidade, nunca existiram. Em muitos casos, esse conteúdo é patrocinado e veiculado nas redes sociais como se fosse parte da comunicação oficial da empresa. A consequência direta é o comprometimento da marca. Colaboradores, clientes ou parceiros podem associar o conteúdo falso à comunicação legítima da organização, gerando confusão e perda de confiança. Mesmo após a contenção do incidente, o esforço para reconstruir credibilidade pode ser longo e custoso.

Vale observar que criminosos não precisam necessariamente usar a face da vítima, especialmente quando há características físicas ou de estilo marcantes que induzem esse tipo de associação.

Vinculação entre ataques direcionados e incidentes de marca

Outro ponto de atenção é a sobreposição entre ataques de engenharia social e abusos de marca. Criminosos têm combinado deepfakes de executivos com domínios falsos, e-mails visualmente legítimos e perfis clonados em redes sociais para criar contextos críveis. O uso de identidade corporativa nesses ataques — logos, nomes, slogans, até elementos de tom de voz institucional — reforça a ilusão de autenticidade.

Para os times de segurança da informação, isso representa um desafio que vai além da detecção de ameaças técnicas. Envolve monitorar sinais de abuso de imagem e narrativa corporativa em múltiplos ambientes e responder rapidamente para evitar que uma ação isolada escale para uma crise institucional.

A resposta a esse tipo de ataque precisa ir além do monitoramento: exige intervenção direta nos ambientes onde os conteúdos fraudulentos estão sendo veiculados. É nesse contexto que entra o conceito de takedown — um processo de remoção de conteúdos ou ativos maliciosos, solicitado junto a plataformas, provedores de serviço, redes sociais, registradores de domínio, entre outros intermediários.

O que pode ser removido através do Takedown

No caso de fraudes com deepfakes, o takedown se torna uma ferramenta essencial quando o conteúdo manipulado viola as diretrizes de uso da plataforma em que o conteúdo está hospedado. Perfis falsos de executivos em redes sociais, vídeos hospedados em sites de terceiros, anúncios enganosos, páginas de phishing com identidade visual da empresa — todos esses elementos são, em regra, passíveis de remoção.

A lei federal Take It Down Act nos Estados Unidos exige que plataformas dedicadas à distribuição de conteúdo criado por usuários tenham um canal específico e ágil para **denunciar e remover em até 48 horas** as representações visuais íntimas, sejam elas autênticas ou geradas por IA, como um deepfake. Essa lei não exige que a imagem seja pornográfica nem retrate o rosto da pessoa – qualquer característica que permita identificar o indivíduo, como uma marca de nascença, pode ser suficiente para justificar o pedido de Takedown nos termos da legislação.

Por outro lado, **nem todo tipo de conteúdo gerado com IA pode ser removido com a mesma facilidade.** Conteúdos que simulam contextos fictícios sem infringir diretamente uma marca registrada, por exemplo, podem escapar dos critérios de moderação tradicionais. Nesses casos, o papel da segurança da informação é avaliar o risco reputacional, buscar evidências técnicas de manipulação e acionar canais alternativos — como comunicação institucional, jurídico e gestão de crise.

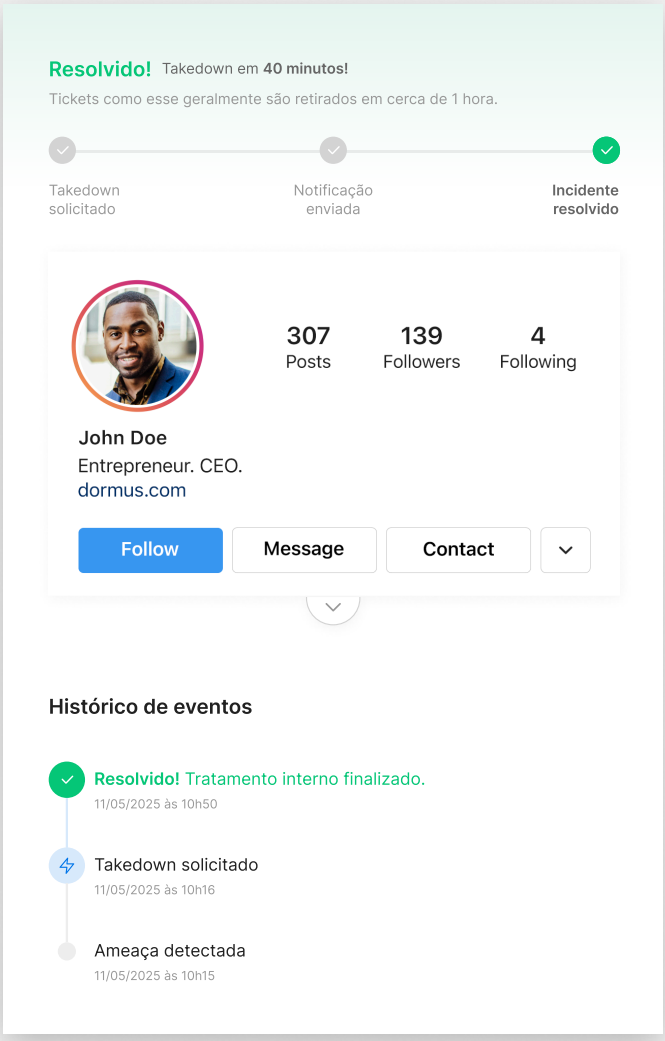
O que pode (ou não) ser removido em uma fraude com deepfake

Elemento utilizado na fraude	Domínio	Passível de takedown?
E-mail em nome do executivo com domínio similar ao oficial	Corporativo	✓ **Sim** O domínio pode ser removido
Perfil falso com nome e foto do executivo no LinkedIn ou Instagram	Pessoal (impacto corporativo)	✓ **Sim** O perfil pode ser removido
Página falsa com identidade visual da empresa (logos, cores, nome comercial)	Corporativo	✓ **Sim** Pode ser removida por violação de marca
Imagem ou vídeo que contenha cena íntima cuja distribuição não foi autorizada, seja ela autêntica ou falsa	Ambos	✓ **Sim** Em conformidade com o Take It Down Act, a remoção deve ser realizada em até 48 horas a partir de 2026.
Vídeo deepfake do executivo promovendo um investimento falso com uso da marca	Ambos	✓ **Sim** Passível de remoção (plataformas e anúncios)
Vídeo deepfake do executivo sem uso de marca, mas com aparência realista	Pessoal	✓ **Sim** Passível de remoção
Vídeo em plataforma estrangeira com simulação genérica de executivo, sem marca	Ambos	× **Não** Difícil remover sem violação explícita
Domínio falso simulando subdomínio oficial (ex: financeiro.suaempresa.online)	Corporativo	✓ **Sim** Pode ser removido com base em abuso de marca
Áudio clonado do executivo enviado por app de mensagem	Pessoal (impacto interno)	⚠ **Não diretamente** Difícil de rastrear ou remover
Vídeo falso publicado em grupo fechado (WhatsApp, Telegram, etc.)	Ambos	⚠ **Não diretamente** Mas os canais podem ser monitorados e investigados
Postagem falsa simulando declaração institucional assinada por executivo	Corporativo	✓ **Sim** Pode ser removida como violação de imagem e marca

A plataforma Axur viabiliza o processo de identificação e solicitação de takedown em escala, com suporte a diferentes tipos de ativos e canais.

Em cenários onde a velocidade de resposta é determinante — como uma campanha de fraude com a imagem do CEO circulando em um perfil falso — o tempo de reação pode ser o fator decisivo entre conter um incidente ou permitir que ele escale.

Para as equipes de segurança, essa capacidade representa não apenas uma reação mais rápida, mas também um ponto de integração com outras áreas da empresa, ampliando a visibilidade e a coordenação em incidentes de impacto externo.



Conclusão

Como vimos ao longo deste material, a superfície de ataque não se limita a sistemas. Ela se estende à identidade institucional e à presença pública dos líderes da organização — que, quando manipuladas, tornam-se ferramentas para golpes com potencial de causar prejuízos financeiros, crises de imagem e impactos estratégicos.

A tendência é de que esses ataques se tornem ainda mais convincentes. Tecnologias de clonagem de voz em tempo real, avatares interativos e a integração entre deepfakes e modelos de linguagem (LLMs) já começam a tornar possível a criação de executivos sintéticos capazes de dialogar com vítimas de forma personalizada e adaptável. A escalabilidade desses ataques exigirá, cada vez mais, uma resposta coordenada entre times de segurança, jurídico, comunicação e governança.

A plataforma Axur foi desenvolvida para apoiar empresas nesse novo cenário. Automatizamos a detecção de ameaças externas, o monitoramento de perfis falsos e a execução de takedown em escala, oferecendo visibilidade e agilidade na contenção de fraudes que envolvem a imagem de executivos e a identidade da marca.

Se sua empresa já está lidando com incidentes desse tipo — ou quer se preparar para enfrentá-los com confiança — estamos à disposição para apoiar sua estratégia de defesa.