A nova era do Phishing

Evolução, tendências e os principais aprendizados de mais de 2 milhões de takedowns bem-sucedidos

AXUR

Sumário executivo

O phishing evoluiu e se tornou um dos ataques mais persistentes e sofisticados da atualidade. Ele deixou de ser apenas um golpe por e-mail e agora explora múltiplos canais, incluindo redes sociais, SMS, anúncios pagos e até deepfakes. Além disso, criminosos evitam mencionar diretamente marcas em domínios e códigos HTML para driblar detecções tradicionais.

Empresas que ainda tratam o phishing apenas como um problema de TI podem estar subestimando seu impacto. Essa ameaça compromete diretamente a reputação da marca e a segurança de clientes e parceiros, gerando custos adicionais com suporte, ressarcimentos e recuperação de confiança.

Principais mudanças em phishing

→ Multicanalidade:

o phishing pode começar por e-mail, mas também com um SMS, redes sociais, anúncios pagos ou chamadas telefônicas.

→ Uso de inteligência artificial:

criminosos criam páginas altamente convincentes e dificultam a detecção por padrões tradicionais.

→ Atenção ao mobile:

sites falsos são otimizados para dispositivos móveis, onde a URL fica menos visível e a interação é mais rápida.

→ Publicidade como vetor:

golpes são promovidos por meio de anúncios online, explorando o interesse recente das vítimas. Ao direcionar os anúncios para quem já pesquisou sobre o tema, criminosos aumentam a chance de engajamento, aproveitando o momento em que a pessoa está mais propensa a clicar em uma "promoção imperdível".

A resposta ao phishing: monitoramento e takedown

O combate ao phishing exige estratégias que vão além do monitoramento passivo. A identificação de ameaças precisa ser acompanhada de uma ação rápida para minimizar danos.

O uso de Inteligência Artificial avançada, como o VLMs (Vision Language Models), permite investigar campanhas de phishing em tempo real e detectar ameaças antes que elas atinjam vítimas.

Após a detecção, é importante entender o processo de takedown, que remove páginas fraudulentas, perfis maliciosos e anúncios enganosos da internet.

Na prática, o phishing não só evolui constantemente, mas também se torna mais difícil de detectar e combater. Estratégias tradicionais já não são suficientes para conter a ameaça.

A seguir, veja como o monitoramento avançado, inteligência artificial e takedown automatizado transformam a resposta ao phishing, garantindo proteção real para marcas e consumidores.

Leia o conteúdo completo →

Phishing: uma ameaça às marcas

O golpe de phishing mais tradicional é aquele que se passa por uma instituição conhecida pela vítima, como um banco, uma varejista ou uma entidade governamental. Nessa perspectiva, o phishing é inseparável das marcas e identidades legítimas das quais os criminosos se valem para aplicar esse golpe.

O phishing definitivamente é um problema para redes corporativas e usuários finais, que recebem uma quantidade imensa de e-mails indesejados. Em 2023, o Google revelou que bloqueia 15 bilhões de mensagens de spam e phishing diariamente.

Contudo, o phishing hoje não se limita mais a uma modalidade de comunicação. Golpes de phishing podem começar por e-mail, por SMS, chamada telefônica ou por publicidade em redes sociais. Aos poucos, a fraude direciona a vítima para outros canais mais adequados ao objetivo de roubar as informações desejadas.

Já é possível afirmar que o phishing é uma fraude "multicanal" ou "multivetorial". Não é mais um problema limitado ao e-mail.

O que não mudou, por outro lado, é que o phishing continua copiando a identidade visual das empresas e se aproveitando das marcas para enganar seus clientes.

Caso ainda não tenham olhado para o phishing como uma ameaça à sua marca, as empresas têm muito a ganhar com essa postura.

Ao tratar o phishing como um problema de marca – da mesma forma que pirataria, ofertas ilegítimas e outros casos semelhantes –, a empresa passa a acompanhar mais de perto como é a experiência online de seus consumidores ou parceiros que podem ser vítimas desse tipo de golpe.

Ao mesmo tempo, é possível adotar medidas de contenção – como a prática de takedown – para diminuir o impacto do phishing e proporcionar uma experiência digital mais limpa e segura. Isso evita que os consumidores caiam em golpes que podem trazer danos indevidos à reputação da marca e até custos decorrentes do atendimento aos clientes que sofreram esses ataques.

Entender que o phishing não é um problema só de quem o recebe é o primeiro passo para que uma empresa possa combater essa ameaça de forma eficaz.

Fundamentos do Phishing

Ainda que o phishing tradicional remeta à ideia de um e-mail falso tentando se passar por uma instituição financeira, esse entendimento restrito da fraude não é adequado ao cenário atual.

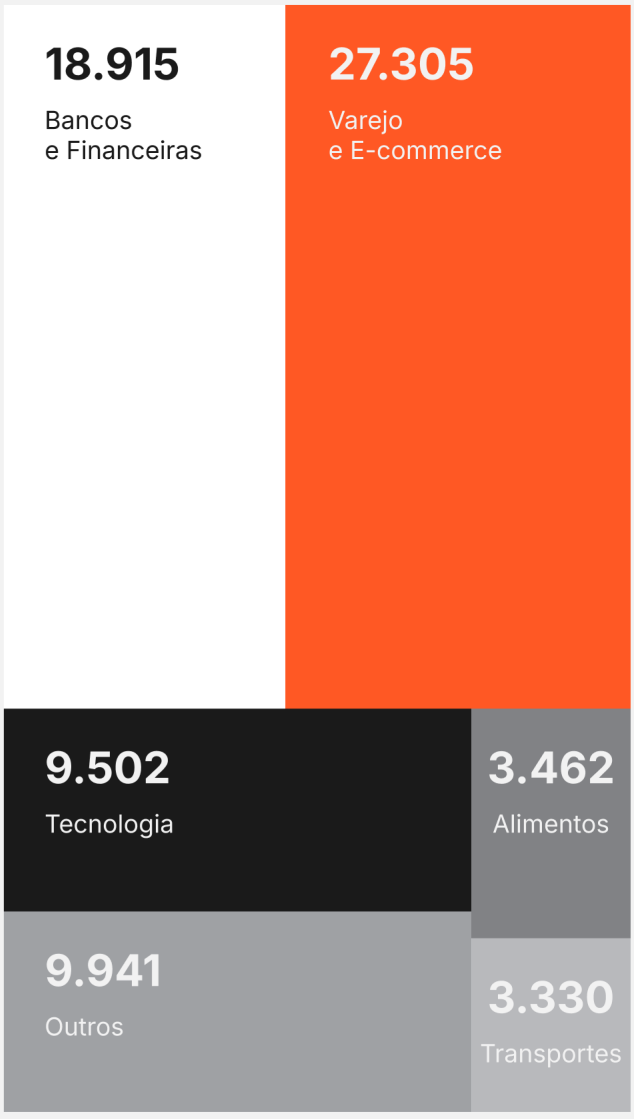
É melhor entender o phishing como todo o conjunto de fraudes em que o criminoso utiliza uma marca de forma indevida e prepara uma narrativa para induzir uma visita a um site ou outra ação que leve ao vazamento de informações.

Sob essa perspectiva, não se deve imaginar que o phishing só possa existir em um canal específico (como o e-mail), nem que o objetivo do phishing é sempre o mesmo (informações financeiras).

O phishing também não se restringe ao segmento B2C. Criminosos se valem da identidade de fornecedores ou parceiros para obter informações e até credenciais corporativas, visando ataques contra redes corporativas para golpes como ransomware ou roubo de dados.

Com a popularização de suítes de escritório em nuvem e plataformas de software como serviço (Software-as-a-Service, SaaS), o phishing pode tentar assumir diversas identidades para roubar credenciais que dão acesso a ativos empresariais, inclusive a partir das mesmas fraudes que atingem usuários finais.

Setores mais atingidos pelo phishing



Setores mais afetados por phishing em 2024, segundo as detecções da plataforma Axur.

Tipos de phishing

O phishing não se restringe mais a nenhum canal ou metodologia.

Relativo ao canal

→ **Smishing:** SMS

→ **Vishing:** Chamada de voz

→ **Quishing:** Códigos QR

→ **Fake Ads/malvertising:** Anúncios publicitários

Relativo à técnica

→ **Spear phishing:** Mensagem elaborada para alvos específicos

→ **Fake News:** Elaboração de conteúdos noticiosos que promovem páginas falsas

→ **Pharming:** Utiliza sequestro de DNS/ domínio

Prejuízos do phishing para as marcas

Para as vítimas diretas do phishing, sejam elas empresas ou consumidores finais, a fraude pode causar imensos transtornos para reaver os valores roubados, cancelar cartões de crédito e trocar senhas.

→ As empresas cujas marcas aparecem nos golpes de phishing também acabam tendo alguns prejuízos materiais ou imateriais.

Insatisfação do cliente que foi vítima do golpe, possivelmente com prejuízos decorrentes desse sentimento.

→ Custos com atendimento às vítimas do phishing, que acreditavam ter lidado com a marca de forma legítima.

→ Vendas ou oportunidades perdidas associadas à sensação de insegurança dos clientes ou possíveis parceiros.

Tecnologias essenciais

→ Detecção de phishing mobile
É preciso simular acessos de dispositivos móveis para garantir que o phishing seja detectado. Muitas páginas fraudulentas são configuradas para abrir apenas em smartphones, impedindo o acesso por notebooks e computadores. Além disso, golpes direcionados a redes sociais e dispositivos móveis frequentemente adotam múltiplas camadas de filtragem, bloqueando coletores menos sofisticados e tornando a detecção por bots ainda mais difícil.

→ IA para análise avançada em escala
Milhões de páginas e artefatos precisam ser analisados diariamente para detectar casos de phishing. A inteligência artificial é capaz de realizar inspeções visuais, detectando marcas e semelhanças de layout e cores com páginas legítimas. Tudo isso acontece em uma fração do tempo que a análise humana levaria, sem comprometer a qualidade da detecção.

O phishing e a cibersegurança externa

Empresas não precisam depender somente das soluções de cibersegurança dos usuários para bloquear ataques de phishing. Com uma abordagem de Cibersegurança Externa para o combate a fraudes, é possível monitorar o uso indevido da marca, identificar ataques de phishing e tomar medidas proativas para mitigar os efeitos das campanhas maliciosas.

O phishing ficou mais sofisticado. E mais difícil de detectar.

Se as técnicas utilizadas pelos golpistas ainda fossem as mesmas de décadas atrás, o phishing dificilmente se manteria como um dos principais cibercrimes do mundo. Portanto, o que vemos na prática é uma variação constante nas táticas.
Os experimentos bem-sucedidos dos criminosos se consolidam em tendências.

Infelizmente, os mecanismos de proteção não podem se concentrar apenas nas técnicas mais novas. Se algum truque antigo voltar a ser eficaz por conta de alguma lacuna nas proteções, os criminosos provavelmente voltarão a utilizar as técnicas "antigas".

Um exemplo disso são os e-mails em que a mensagem de phishing é colocada em um anexo. A ideia de esconder os elementos suspeitos do e-mail em um formato alternativo (como PDF, DOC ou XLS) para escapar do anti-spam foi muito utilizada em alguns anos da década de 2010. Depois de uma pausa, esse tipo de mensagem voltou a ter eficácia.

Em outras palavras, alguns truques antigos podem ser "reciclados" quando a mudança na atuação dos filtros – muitas vezes para combater outras fraudes mais novas – acaba reabrindo uma lacuna antiga.

Dito isso, ainda temos algumas tendências relevantes – e realmente mais recentes – que estão marcando os golpes de phishing: a ausência de menções diretas às marcas, o foco no mobile, a publicidade como canal de divulgação e o uso de páginas intermediárias para esconder a finalidade do golpe.

Essas tendências não estão isoladas. Ao contrário, é comum que um mesmo phishing adote uma ou mais dessas estratégias simultaneamente.

Menos menções diretas às marcas

Como o phishing tenta se passar por uma empresa ou entidade conhecida, era possível identificar a maioria dos ataques phishing através da análise de novos domínios registrados e do endereço das páginas.

Em muitos casos, isso permitirá antever uma campanha de phishing que ainda seria lançada pelos criminosos, já que havia um intervalo entre o registro do domínio e o início da campanha para divulgar o site falso.

Porcentagem de domínios

70% dos domínios maliciosos evitam palavras-chave de marcas.

Menção no código HTML

18% não mencionam a marca nem no HTML.

Para evitar que os golpes sejam detectados por essa abordagem, os criminosos estão evitando a inclusão de palavras-chave vinculadas às marcas no domínio.

Isso não significa que o phishing deixou de utilizar as marcas para fraudar os consumidores. Eles apenas mudaram a forma de criar a associação com a marca para que ela seja menos evidente para algoritmos tradicionais.

Nos casos mais simples, a palavra-chave da marca é simplesmente deslocada para elementos como os subdomínios.

Devido ao foco nos ataques mobile, a ausência da marca no domínio não traz um prejuízo relevante para a eficácia da fraude. O tamanho reduzido da barra de endereço do navegador do smartphone dificulta a visualização completa da URL, dando mais flexibilidade para a manipulação do endereço.

Para retirar qualquer menção textual à marca, os criminosos podem apostar em referências visuais, seja através da inclusão do logotipo ou da utilização de elementos que remetem à identidade visual ou à empresa, como produtos e imagens de fachadas de lojas.

A inteligência artificial é uma das melhores ferramentas para enfrentar esse desafio. Através da aprendizagem de máquina, IA identifica essas semelhanças e associa os elementos visuais de forma muito parecida com um humano, podem vincular um site de phishing a uma marca que não foi mencionada.

Phishing multicanal com foco em mobile

Os dispositivos móveis sempre foram muito convenientes para a comunicação, mas as limitações na funcionalidade criavam um obstáculo para a ideia de que uma fraude poderia ocorrer totalmente no ambiente mobile. Os smartphones mudaram isso.

Além disso, as pessoas hoje simplesmente passam mais tempo nos dispositivos móveis. De acordo com dados da Statcounter, 61% de todo o tráfego web é hoje proveniente de smartphones.

Como resultado, muitos golpes de phishing já têm dispositivos mobile como alvos prioritários.

O mobile como alvo prioritário

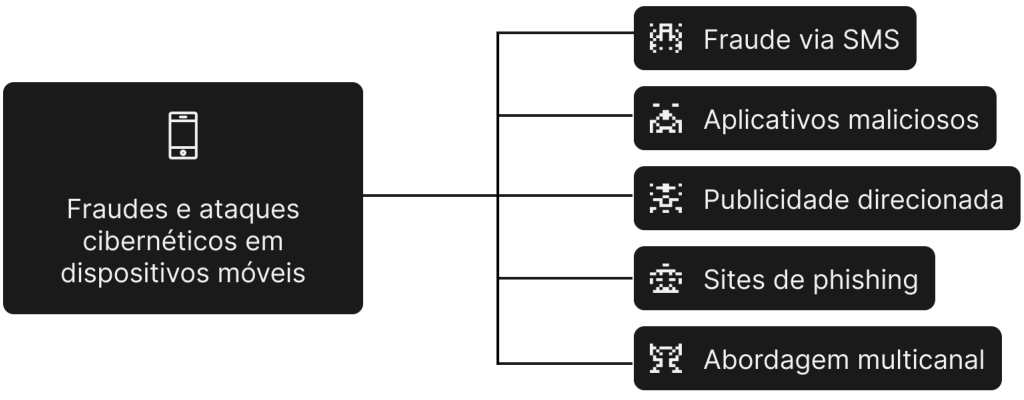
→ Promoção de fraudes por SMS ou outros canais disponíveis apenas no ambiente mobile

→ Divulgação de apps maliciosos

→ Uso de publicidade dirigida em redes sociais e sites de busca para que o phishing seja promovido apenas para usuários em dispositivos móveis

→ Implementação de filtros nos sites de phishing que redirecionam acessos de outros dispositivos (como notebooks) para páginas legítimas, dificultando a análise da página

Criminosos podem começar a fraude enviando um SMS que divulga um link, um número de telefone ou um aplicativo malicioso, criando um phishing multicanal.



A conectividade ampla dos dispositivos móveis cria muitas oportunidades para os golpistas. A resposta mais adequada é uma abordagem igualmente abrangente para o combate ao phishing, com capacidade para detectar e mitigar todas as modalidades da fraude, seja ela através de SMS, e-mail, publicidade ou aplicativos falsos que podem estar em qualquer loja de aplicativos.

Como garantir a visibilidade sobre o phishing

Como os criminosos utilizam filtros para que as páginas de phishing fiquem menos expostas aos sistemas de segurança, é preciso adotar soluções que contornem essas medidas de contrainteligência do crime.

1. Acesso com sistemas de proxy em múltiplas regiões.

2. Simulação das características técnicas de dispositivos móveis, incluindo capacidades gráficas e tela sensível ao toque.

3. Tentativas de coleta do phishing a partir de outras categorias de dispositivo.

4. Captura de imagem de tela (screenshot) do phishing no formato em que o acesso foi bem-sucedido.

A publicidade como canal de phishing

A prática de promover conteúdos maliciosos utilizando anúncios publicitários não é nova e já foi denominada de malvertising, um termo faz referência à presença de códigos maliciosos (normalmente Javascript).

Os códigos maliciosos no anúncio fazem com que ele seja aberto mesmo que o visitante não interaja com ele, ou até executam outras atividades (como mineração de criptomoeda). O nome "malvertising" também remete à ideia de que o anúncio pode promover malware.

O que está sendo observado hoje, porém, nem sempre se encaixa neste conceito. Alguns dos anúncios não possuem nenhum código malicioso, sendo divulgados em plataformas de redes sociais e em links patrocinados de pesquisas web — locais que restringem bastante o formato e as capacidades técnicas da publicidade veiculada.

Como o anúncio precisa convencer o visitante a interagir, a estratégia mais uma vez remete ao phishing e ao uso indevido de marcas.

O dano decorrente do uso indevido da marca em anúncios tende a ser maior do que o do phishing tradicional. Além de prejudicar os consumidores, essa prática pode criar uma concorrência com a marca em determinados segmentos, como nas palavras-chave de sites de busca. A presença recorrente dessas fraudes pode deixar o consumidor mais relutante.

Nem toda a publicidade falsa se resume a uma tentativa de phishing para roubo de credenciais. Os criminosos também podem tentar promover aplicativos falsos ou realizar fraudes de pagamento (oferecendo promoções ou descontos que não existem) para roubar o dinheiro da vítima ou dados de cartões de crédito.

De uma forma ou de outra, a lógica do golpe quanto ao uso não autorizado de uma marca para atrair o interesse da vítima é exatamente a mesma do phishing. Além disso, o roubo de credenciais também é uma consequência recorrente, seja de forma indireta ou de forma indireta através do aplicativo falso preparado para a fraude.

Por esses motivos, não é muito proveitoso tentar diferenciar as categorias específicas de cada tipo de fraude. Na prática, todas as fraudes são semelhantes e podem ser mitigadas com as mesmas medidas: monitoramento, inteligência em ameaças e takedown.

Como detectar phishing em anúncios?

Anúncios online podem ser altamente dirigidos para determinados públicos, especialmente porque as plataformas os direcionam conforme os interesses mais recentes, como produtos pesquisados. Isso exige que o monitoramento da marca seja realizado através de canais específicos, que dão acesso mais amplo às peças veiculadas.

O que muda no takedown de anúncios?

Em anúncios de plataformas de mídias sociais, a publicidade precisa estar atrelada a uma conta de gerenciamento, que é um perfil da própria rede social. Isso significa que há ao menos 3 pontos para serem notificados no takedown:

Anúncio ativo	**Perfil vinculado**	**Site removido**
o anúncio que está sendo promovido	O perfil da rede social associado ao anúncio	O site que é promovido pelo anúncio

Para garantir a proteção da marca, o Takedown da Axur notifica todas as partes envolvidas no fluxo do golpe através dos anúncios.

Notícias falsas

As notícias falsas e as campanhas de desinformação (conhecidas como "fake news") são um desafio para toda a sociedade. Contudo, assim como os idealizadores dessas campanhas buscam se aproveitar da credibilidade dos meios de comunicação, os golpistas de phishing há muito tempo já vêm se aproveitando da credibilidade das marcas.

Por esta perspectiva, talvez não seja muito surpreendente que os golpistas por trás do phishing tenham percebido que eles também podiam integrar o conceito de fake news em suas fraudes.

O modus operandi normalmente envolve a criação ou manipulação de fatos noticiosos que possam vincular a marca da empresa atacada e publicar a "reportagem" em um site falso de notícias – em geral, copiando as cores e elementos visuais de sites jornalísticos conhecidos.

Essa tática tem 3 vantagens para os criminosos.

→ Oculta o golpe
Quando o phishing é divulgado por meio de anúncios, o aspecto noticioso da página intermediária ajuda a ocultar o golpe e diminui as chances de o anúncio ser bloqueado pela própria plataforma.

→ Abre mais espaço para a narrativa
O conteúdo aparentemente noticioso permite que os golpistas elaborem narrativas mais complexas: promoções que não existem, liquidação por recuperações judiciais e leilões são alguns exemplos que podem estimular uma reação imediata da vítima. Os consumidores estão menos preparados para lidar com essas narrativas do que com outras técnicas antigas do phishing, como o recadastramento da senha.

→ Aumenta a credibilidade
A partir do contato com a notícia forjada que teria sido publicada em um portal confiável, a vítima do phishing fica mais propensa a acreditar na fase seguinte da narrativa da fraude, mesmo que ela não faça menções diretas às marcas.

As notícias falsas também são um veículo poderoso para a introdução de deep fakes produzidos com inteligência artificial no phishing.

Para dar ainda mais credibilidade ao golpe, podem ser usadas falas ilegítimas de executivos ou até vídeos manipulados.

Considerando que as possibilidades abertas por essa estratégia são muitas, é importante que as fraudes sejam acompanhadas de perto por equipes de inteligência em ameaças.

Como um phishing pode utilizar todas essas técnicas?

1. O criminoso cria uma página imitando um site de notícias com uma reportagem sobre a empresa Apple.

2. A página de notícias possui um link para outra página ou aplicativo que tentará roubar dados de clientes da empresa Apple. Como a vítima já veio da página de notícias, a página ou aplicativo que rouba informações não precisa mais de menções diretas à marca.

3. O criminoso prepara peças publicitárias e mensagens SMS para divulgar o site com a notícia falsa. Como todas as visitas esperadas no golpe devem vir de dispositivos móveis, a página pode adotar filtros que só aceitem esse tipo de acesso, bloqueando sistemas de coleta menos sofisticados.

Como identificar mesmo os golpes de phishing mais avançados?

O phishing é um golpe muito comum e, por isso, impacta diversos negócios. Para quem ainda não possui uma solução de monitoramento, o resultado pode ser surpreendente. A busca por palavras-chave e expressões próprias da marca ou dos produtos pode até ser suficiente para detectar e combater um número significativo de fraudes. Porém, nenhuma empresa é igual. As organizações precisam ter a flexibilidade de coletar informações e utilizar o poder das ferramentas de monitoramento para descobrir a melhor forma de ajustar os processos e políticas que enfrentam essas ameaças.

Como LLMs estão redefinindo a detecção de ameaças

A detecção de ameaças digitais evoluiu significativamente nos últimos anos, e os Large Language Models (LLMs) estão desempenhando um papel central nessa mudança.

Tradicionalmente, a identificação de phishing e fraudes dependia de regras pré-definidas e análise baseada em palavras-chave, um método eficaz, mas limitado diante da evolução das táticas dos criminosos.

Com a aplicação de Vision Language Models (VLMs), é possível analisar não apenas o conteúdo textual de um site malicioso, mas também sua estrutura visual e semântica. Esse modelo permite detectar padrões avançados de personificação, identificar sinais de fraude em páginas que evitam menções diretas a marcas e extrair insights que antes dependiam exclusivamente de análise humana.

Ao processar grandes volumes de dados diariamente, os VLMs aumentam a precisão da detecção, reduzindo falsos positivos e identificando ameaças que passariam despercebidas por sistemas tradicionais. Além disso, o uso de inteligência artificial generativa permite a descrição detalhada de páginas suspeitas, fornecendo contexto e facilitando investigações.

Aplicação na detecção de phishing e fraudes

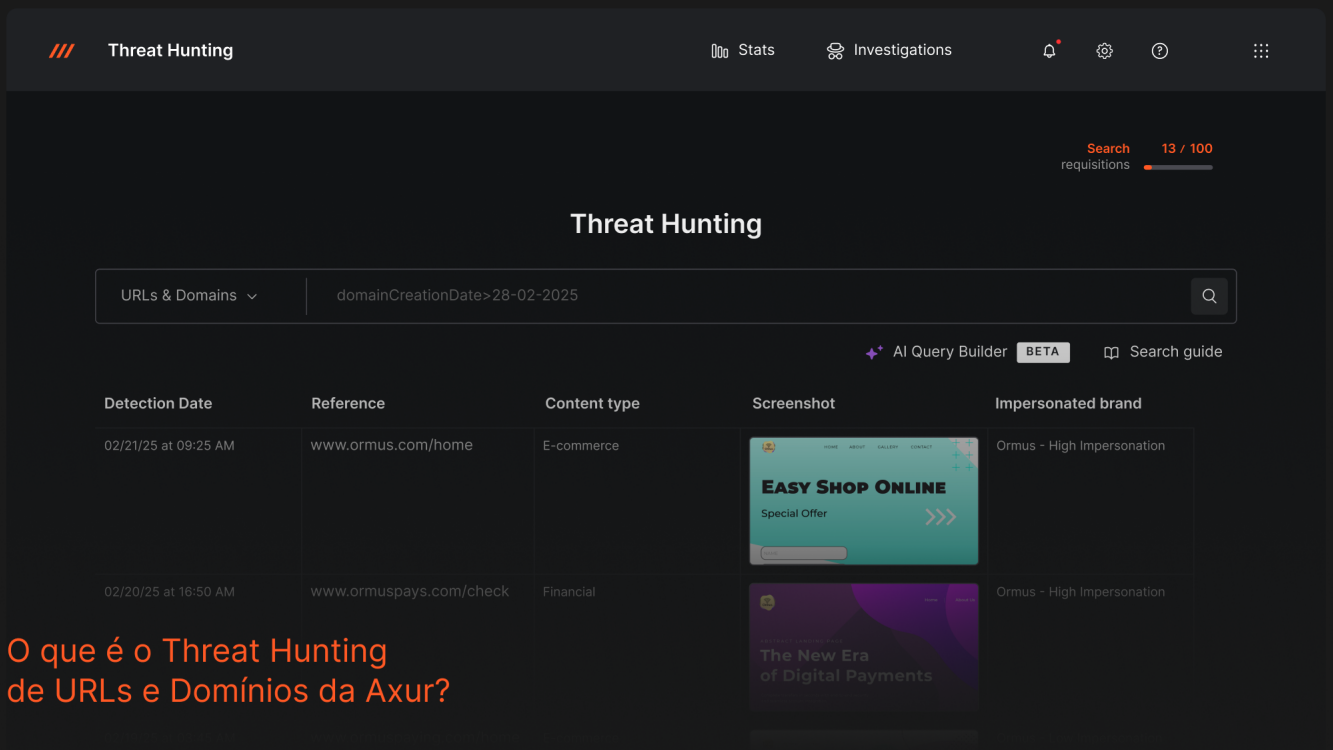
A Axur incorporou essas capacidades no Clair VLM (Cyber Lens for Anomaly and Impersonation Recognition), um modelo proprietário baseado em VLMs e treinado com mais de 15 anos de dados sobre ameaças digitais. Esse modelo inspeciona mais milhões de sites por dia, analisando tanto os elementos visuais quanto os dados estruturais das páginas.

O Clair identifica tentativas de personificação de marcas, verifica solicitações de credenciais, pagamentos ou senhas e gera descrições detalhadas das páginas analisadas.

Esse processo ocorre de forma totalmente automatizada, garantindo eficiência na detecção e minimizando a necessidade de intervenção manual.

Diferente de abordagens tradicionais que dependem apenas de palavras-chave ou bloqueios baseados em listas de domínios, a integração entre VLMs e análise de ameaças permite uma visão mais ampla e contextualizada dos ataques, oferecendo maior precisão e cobertura na proteção contra fraudes digitais.

A detecção avançada de Phishing & Domínios da Axur reúne informações coletadas e sinais extraídos por algoritmos e por inteligência artificial para identificar as fraudes mais complexas e inovadoras.



O que é o Threat Hunting de URLs e Domínios da Axur?

O Threat Hunting é uma ferramenta de descoberta e inteligência. Ele oferece a capacidade necessária para ampliar a visibilidade sobre as fraudes sem impactar os processos já estabelecidos e o tratamento automatizado de incidentes.

Para quem o utiliza, o Threat Hunting funciona como um serviço de busca em que os resultados são informações relativas a fraudes digitais e outros sinais de inteligência em ciberameaças.

Uma das categorias do Threat Hunting é a de URLs e Domínios, é ideal para investigar campanhas de phishing. Ela permite localizar páginas suspeitas a partir de critérios como endereço, data e marcas identificadas.

Nos resultados, o analista tem acesso ao endereço do site, à captura de tela, ao endereço IP, ao código HTML da página e a mais de duas dezenas de sinais coletados.

Inteligência artificial em escala

Uma das principais vantagens da detecção de Phishing & Domínios da Axur é a capacidade de análise avançada impulsionada pelo modelo Clair VLM. Em vez de depender apenas de critérios objetivos, como a data de registro de um domínio ou sua primeira detecção, o Clair processa e interpreta atributos críticos para identificar possíveis fraudes:

Idioma AI

Detecta os idiomas presentes na página, permitindo priorizar investigações conforme o público-alvo do golpe.

Tipo de conteúdo AI

Identifica se a página se assemelha a telas de login, formulários de pagamento ou e-commerce, mesmo sem menções explícitas a marcas.

Solicitação de credenciais AI

Verifica se há campos para inserção de e-mails, usuários e outras informações sigilosas.

Solicitação de senha AI

Avalia a presença de campos específicos para login e autenticação.

Solicitação de pagamento AI

Examina elementos textuais e estruturais para identificar páginas que tentam capturar dados financeiros.

Ao combinar essas análises com o aprendizado de máquina, o modelo permite uma detecção mais precisa e contextualizada, garantindo que analistas encontrem fraudes que poderiam passar despercebidas por métodos convencionais.

Em todos os sinais processados pelo modelo Clair, é possível conferir exatamente qual foi a resposta para cada critério.

O modelo também identifica as marcas presentes na página e mensura o nível de semelhança com a identidade visual de cada uma das marcas identificadas.

Combinando esses critérios, você pode localizar páginas que tenham sido criadas recentemente, que tem alta semelhança com uma marca especificada e que possuam um dos outros critérios (solicitação de pagamento, senha ou credencial).

Os critérios técnicos (como códigos HTTP, endereço IP e outros) também podem ser usados na pesquisa para desvendar campanhas de phishing que estejam compartilhando de uma mesma infraestrutura.

Com essa visão, analistas podem se aprofundar em cada campanha de phishing, indo além do que foi detectado pelos sistemas automatizados. Por se tratar de uma busca imediata, é possível experimentar vários critérios.

O que for descoberto pode informar a tomada de decisões e o aprimoramento dos critérios de automatização e combate à fraude.

Identificar o phishing é apenas o primeiro passo. Depois de detectar uma ameaça, é fundamental agir rapidamente para impedir que vítimas caiam no golpe e minimizar os impactos para a marca. É aqui que entra o takedown, removendo do ar páginas fraudulentas e perfis maliciosos antes que causem mais danos.

Como o takedown combate o phishing e outras fraudes digitais

O takedown é a ação de derrubar um conteúdo que está disponível na web. Se o conteúdo for um perfil em uma rede social, ele será suspenso após um takedown. Se for um site ou domínio, o takedown ocorre quando o provedor de hospedagem ou a entidade registradora cancelam o serviço, fazendo com que a página seja derrubada.

O uso de marca sem autorização pode ser considerada uma atividade ilegal. Para esses casos, o tipo de denúncia mais recorrente é a alegação de violação de direito de propriedade intelectual para retirada de perfis que se apropriam de sinais distintivos da marca, como nome e logo, em benefício próprio.

A DMCA (Digital Millennium Copyright Act, Lei dos Direitos Autorais do Milênio Digital) é uma lei de direitos autorais dos Estados Unidos que permite aos provedores digitais a isenção de responsabilidade por violação de direitos autorais, caso eles removam prontamente o conteúdo

Os provedores de serviços de internet (ISPs) também possuem políticas que definem regras e limites para a utilização do ambiente disponibilizado aos usuários ou clientes.

Como o phishing e muitas outras fraudes digitais violam essas políticas, os provedores não permitem que elas fiquem no ar após uma notificação que os deixe cientes sobre a existência do problema.

Quando uma empresa monitora sua marca e identifica fraudes na web, enviar notificações de Takedown é uma evolução natural da estratégia de combate a esses golpes. Quando o conteúdo malicioso é derrubado, a ameaça é neutralizada, evitando que os clientes associem as narrativas fraudulentas à marca.

Uma estratégia assídua de takedown também normalmente desmotiva os criminosos a utilizar uma determinada marca. Afinal, os criminosos perdem horas do seu "trabalho" quando uma fraude é derrubada. Por este motivo, eles preferem abusar de marcas que não enviam notificações de takedown, uma vez que estas fraudes ficarão mais tempo no ar.

Ainda assim, é preciso considerar alguns aspectos técnicos do takedown para que ele seja realizado corretamente e estar preparado para entender como devemos medir a eficácia desse método.

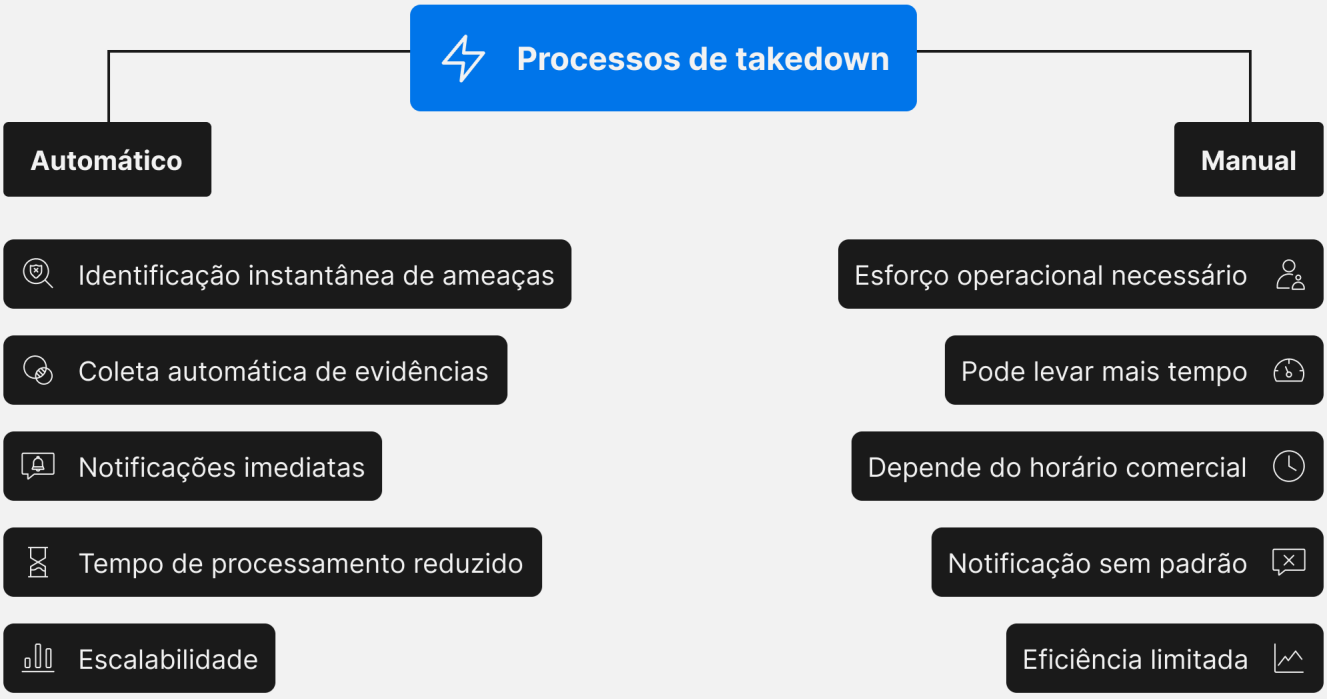
Aspectos técnicos do takedown

O takedown pode ser solicitado por qualquer pessoa e não tem custos. Em geral, basta enviar um e-mail. Contudo, essa simplicidade é um tanto ilusória, uma vez que diversos fatores influenciam a eficácia do pedido. Quando feita incorretamente, a solicitação de Takedown pode demorar a ser atendida ou até ser ignorada, mesmo que o conteúdo denunciado esteja irregular.

O primeiro passo é identificar a violação do ponto de vista do provedor. Para a marca, pode ser evidente que uma publicação ou página é uma fraude. Porém, o provedor terá de enquadrar o incidente como uma violação de suas políticas de uso para justificar a remoção do conteúdo.

É preciso coletar evidências para justificar a notificação de takedown, assegurando o provedor ou a plataforma de que aquele conteúdo não está de acordo com suas diretrizes.

As evidências coletadas precisam ser comunicadas por meio do canal correto, utilizado uma linguagem apropriada (tanto em termos do idioma da solicitação como dos termos técnicos empregados) e um tom adequado. Os Takedowns automatizados têm mais chances de serem acatados mais rapidamente pelas organizações.



Em casos de phishing, a Axur também denuncia a página às instituições que mantêm filtros de navegação segura (Web Safe Reporting), o que vai gerar uma alerta nos navegadores dos usuários, mesmo que a página ainda não tenha sido retirada do ar pelo provedor. Essa ação ajuda a reduzir o contato das possíveis vítimas com o conteúdo malicioso.

Como funciona o Web Safe Reporting

Identificação de Phishing

Axur identifica uma tentativa de phishing e notifica para takedown

Notificação de Web Safe

Axur reporta a página de phishing às instituições

Geração de alerta

As instituições geram um alerta para os navegadores

Usuários alertados

Usuários são alertados sobre a ameaça de phishing enquanto a página não caiu

De todo modo, os criminosos quase sempre precisam envolver algum provedor mais conhecido e respeitável em uma fraude digital. Por exemplo, é muito difícil fazer com que uma fraude chegue a um grande público sem utilizar um serviço de e-mail respeitável e uma rede social conhecida. Se os criminosos enviam um e-mail de phishing a partir de um provedor qualquer, é bastante provável que a mensagem seja classificada como spam.

Por isso, quase sempre há uma preferência por provedores de maior renome, que também são aqueles que atendem prontamente às solicitações de takedown — desde que elas sejam feitas corretamente.

A Axur é reconhecida como uma entidade confiável e tem acesso a canais que agilizam o processo de takedown em diversos provedores.

Veja na prática: como a automação reduz o tempo de exposição da ameaça

A tecnologia da Axur identificou um phishing que replicava o layout de um cliente e gerou um ticket automaticamente às 22h47. Com a automação do sistema:

→ Às 22h55, o takedown foi solicitado automaticamente, e a ameaça foi movida para tratamento.

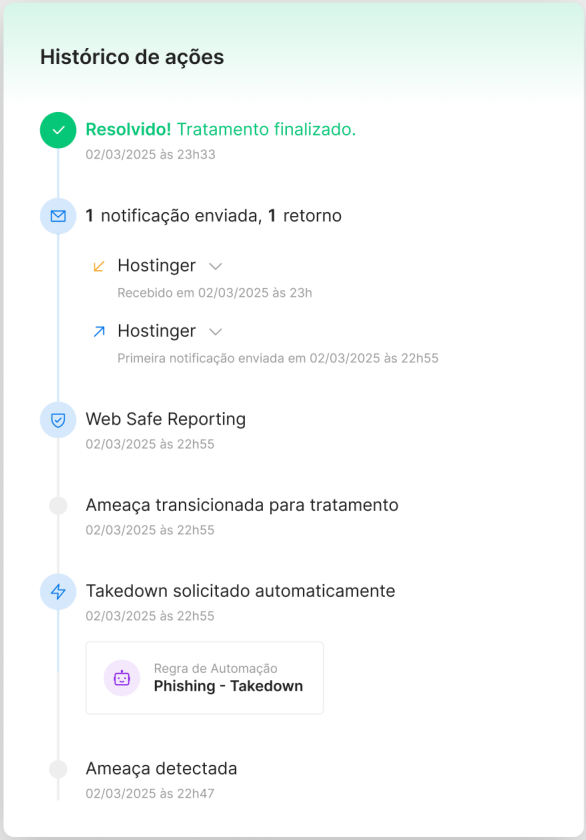
→ No mesmo horário, 15 entidades parceiras foram notificadas via Web Safe Reporting, reduzindo a exposição e o acesso à fraude.

→ Às 22h55, a primeira notificação foi enviada ao ISP, reforçada por uma segunda notificação logo em seguida.

→ Às 23h00, a resposta foi recebida, confirmando o andamento da remoção.

→ Às 23h33, o domínio foi confirmado como inacessível, finalizando o tratamento.

Com essa abordagem, o phishing foi detectado, analisado e removido em apenas 46 minutos, minimizando o risco de vítimas acessarem a página maliciosa.



O que pode ser derrubado pelo takedown

- Hospedagem de páginas que violam marcas registradas (para phishing ou SEO)
- Distribuição de malware (incluindo trojans disfarçados de software legítimo)
- Criação de perfis ou contas de mensagem que usam a identidade de terceiros (incluindo marcas), sem indicar que é uma paródia
- Envio de e-mails em massa não autorizados pelo destinatário (spam)
- Veiculação de anúncios que utilizam imagens ou textos enganosos, fazendo usuários acreditarem que foram feitos por outra marca ou pessoa
- Atividade de robôs ou campanhas que impulsionem conteúdo de maneira inorgânica
- Armazenamento ou distribuição de dados corporativos roubados ou informações privadas
- Pirataria e dados protegidos por direito autoral

O que não pode ser derrubado pelo Takedown

→ Avaliações ou opiniões negativas sobre a marca

Se houver indícios do envolvimento de robôs ou impulsionamento artificial, as publicações inorgânicas podem ser denunciadas

→ Informações falsas ou acusações que não caracterizem uma violação das diretrizes das plataformas

Muitas plataformas são bastante específicas a respeito dos limites envolvendo "fake news". A regra tende a ser mais rigorosa em peças de publicidade ou conteúdo impulsionado do que em publicações de usuários

→ Domínios estacionados ou sem conteúdo

Domínios com nomes similares ao da marca, mas sem conteúdo hospedado, não podem ser derrubados por takedown. Entretanto, monitorar esses domínios para solicitar a remoção caso eles comecem a ser usados para fins maliciosos pode ser o caminho.

Métricas de sucesso em takedown

Uma das vantagens do Takedown como estratégia de combate a fraudes é que existem métricas confiáveis para comprovar sua eficácia e acompanhar o trabalho que é realizado.

Na Axur, todo o processo de Takedown é transparente e pode ser acompanhado em nossa plataforma, etapa por etapa.

% Taxa de sucesso

A taxa de sucesso correspondente à proporção das solicitações que resultam em um Takedown por parte do provedor.

Um takedown bem-sucedido remove o conteúdo do ar, neutralizando a ameaça.

A taxa de sucesso da Axur é de 98,9%. Na prática, quase todas as solicitações são atendidas.

Primeira notificação

Para acelerar o Takedown, é preciso que o provedor seja notificado o quanto antes.

O tempo para o envio da primeira notificação demonstra a capacidade para coletar evidências e acionar o canal correto para dar início ao processo de remoção do conteúdo.

A Axur envia a primeira notificação em uma **mediana de 4 minutos** para quase todos os tipos de incidentes.

Uptime

O uptime afere o tempo que as fraudes notificadas normalmente permanecem no ar até serem removidas pelos provedores. Esse tempo pode variar bastante de um provedor para o outro e leva em conta medidas de reforço, como a repetição da notificação.

A mediana de uptime das **fraudes notificadas pela Axur é de 9 horas**, o que significa que elas tendem a ficar poucas horas no ar.

Duração

Em algumas situações, as fraudes podem ser restabelecidas por criminosos que não querem perder o esforço empenhado na propagação do golpe. Portanto, deve-se continuar monitorando a fraude e refazer a notificação caso o conteúdo volte ao ar, para que o Takedown tenha um efeito duradouro.

A Axur garante que o conteúdo removido permanecerá offline por 15 dias após o takedown. Novas notificações neste período são realizadas sem custo adicional.

Viabilidade e custo-benefício do takedown

Um dos maiores desafios do takedown é a escala. Devido ao volume elevado de golpes na web, uma empresa pode ter dificuldade para notificar todos os incidentes, deixando uma parcela considerável das fraudes fora dessa estratégia.

A chave para solucionar o desafio da escala é a automação. Na Axur, 86% dos takedowns são totalmente automatizados. É possível definir critérios para iniciar uma solicitação de takedown automaticamente, ou escolher incidentes específicos e iniciar o processo com um único clique. Depois disso, basta aguardar até a fraude ser retirada do ar.

A automação da Axur incorpora fluxos para 400 ISPs de todo o mundo, entre provedores de hospedagem, entidades de registro de domínio e plataformas de redes sociais. Isso também garante o envio de múltiplas notificações para incidentes com vários pontos de contato com a fraude (perfil e publicação em rede social, página web, anúncio, por exemplo), de modo que todos possam ser derrubados e a fraude seja completamente neutralizada.

Com parâmetros auxiliados por inteligência artificial, é possível definir um takedown automático para casos de avaliação de risco elevada, quando há menção direta à marca ou em circunstâncias específicas (quando um perfil tem um determinado número de seguidores, por exemplo). A marca permanece protegida 24 horas por dia, 7 dias por semana.

Um aspecto importante é que, como as métricas do takedown são muito claras, é fácil verificar o retorno sobre o investimento nessa abordagem em termos de fraudes derrubadas. A Axur cobra apenas pelos takedowns bem-sucedidos.

Assim, o takedown é uma abordagem de combate a fraudes muito completa e eficaz, especialmente quando estiver aliado a métricas transparentes e uma tecnologia de automação que permita sua aplicação sistemática e integral.

5 mitos sobre takedown que talvez você tenha ouvido

1. Temos um SLA de garantia de remoção

Realidade: nenhuma empresa pode garantir a remoção de conteúdos maliciosos dentro de um prazo fixo, porque a remoção depende de terceiros (provedores, redes sociais, hosts, registrars). O que pode ser garantido é um SLA de resposta—ou seja, em quanto tempo a empresa inicia o processo.

2. Removemos 100% dos conteúdos solicitados

Realidade: nem todos os conteúdos podem ser removidos. Alguns provedores ignoram pedidos, alegam liberdade de expressão ou exigem ações legais. O verdadeiro diferencial é a capacidade de escalar e insistir com alternativas (contatos diretos, parcerias, argumentação legal).

3. Removemos qualquer domínio malicioso

Realidade: domínios fraudulentos podem estar em registrars resistentes a pedidos ou protegidos por políticas locais. A remoção pode ser impossível sem intervenção legal ou longos processos administrativos.

4. Não precisamos de evidências adicionais para solicitar remoção

Realidade: a maioria dos provedores exige provas concretas, como prints, logs e links específicos. Simplesmente alegar que algo é phishing sem evidências pode resultar em rejeição ou atraso.

5. Detectamos todas as ameaças antes que causem danos

Realidade: nenhuma empresa pode garantir detecção 100% antecipada. Algumas ameaças só se tornam visíveis após serem ativadas. O diferencial real é a velocidade de resposta e a precisão na análise.

Como o takedown protege sua marca

O takedown é a denúncia de condutas indevidas ou maliciosas aos provedores de nuvem, hospedagem e plataformas de redes sociais.

A denúncia pode derrubar esse conteúdo. Com o takedown, as empresas, entidades de governo e instituições financeiras retomam o controle sobre o uso de suas marcas na Internet.

Usos indevidos de marca e reproduções não autorizadas, como o phishing, podem ser combatidos através do takedown, seja qual for o canal da fraude.

A Axur tem o Takedown mais eficiente do mundo.

O Takedown da Axur

98,9% de taxa de sucesso

Páginas derrubadas após a denúncia

86% Takedowns automatizados

Compatível com 400 provedores e plataformas

9 horas de Uptime

Mediana de uptime até a derrubada

500.000 Takedowns por ano

Escala para qualquer marca

Garantia Axur

15 dias

O monitoramento continua após a queda para garantir que a fraude se mantenha neutralizada.

4 minutos

Mediana de primeira notificação, incluindo em casos de phishing, reduzindo o alcanceda fraude.

Phishing, fraudes digitais e vazamentos de dados

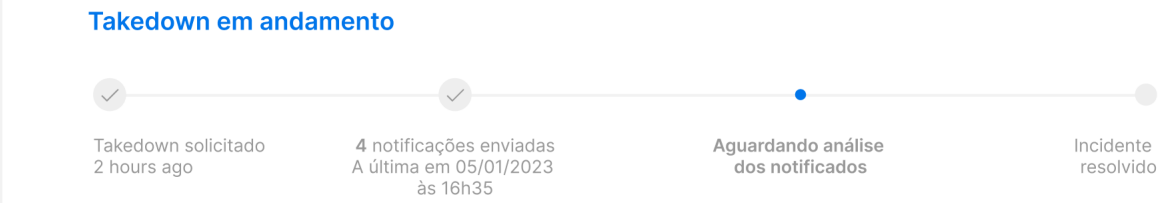
O Takedown combate diversos incidentes que ligam a marca a condutas impróprias.

- Páginas falsas
- Aplicativos ilegítimos
- Perfis falsos em redes sociais
- E mais...

Você no controle

Acesse artefatos e evidências coletadas para o Takedown

Domínio/whois
E-mails
Capturas de tela
Código-fonte



Proteja seus consumidores

30 entidades de segurança notificadas

Além do provedor que abriga a fraude, nosso mecanismo de Web Safe Reporting encaminha denúncias para filtros de phishing que protegem o usuário.

Experimente o melhor Takedown do mundo

FAÇA UMA DEMO

ISO/IEC 27001 Information Security Management CERTIFIED

CleanDNS Trusted Reporter

Descubra todas as nossas soluções em **axur.com**

AXUR