Título: O uso da inteligência artificial na prevenção de fraudes praticadas por falsos advogados em aplicativos de mensagem

Resumo: O artigo examina o uso de inteligência artificial na prevenção de fraudes jurídicas, com foco no golpe do falso advogado em aplicativos de mensagens. Propõe soluções tecnológicas para mitigar o problema, destacando o papel das procuradorias estaduais na formulação de respostas institucionais e na proteção dos jurisdicionados mais vulneráveis.

Palavras-chave: inteligência artificial; golpe do falso advogado; advocacia pública; segurança digital; prevenção de fraudes; aplicativos de mensagens

1. INTRODUÇÃO

A disseminação de golpes praticados por meio do aplicativo WhatsApp tem assumido proporções alarmantes no Brasil, com impactos que extrapolam a esfera individual e alcançam a Administração Pública. Diante da atuação massiva das Procuradorias dos Estados, especialmente nos processos de execução, precatórios e demandas repetitivas, essa realidade impõe desafios adicionais às procuradorias estaduais, que, na linha de frente da defesa do interesse público, precisam lidar com fraudes que comprometem a segurança jurídica, a credibilidade institucional e a eficiência da atuação estatal.

A combinação entre sistemas frágeis de verificação, a desinformação de vítimas — frequentemente em situação de vulnerabilidade — e a promessa de vantagens financeiras imediatas cria um ambiente propício à proliferação de fraudes digitais. Dentre os esquemas que mais têm crescido, destaca-se o golpe do falso advogado, em que criminosos utilizam informações verossímeis e linguagem jurídica convincente para persuadir a vítima da existência de uma ação judicial supostamente vitoriosa, condicionando o recebimento de valores ao pagamento prévio de "custas processuais" inexistentes.

Esse tipo de golpe atinge públicos diretamente vinculados à atuação das Procuradorias Estaduais — como aposentados, servidores públicos, usuários do sistema de saúde e contribuintes em dívida ativa, frequentemente partes em demandas representadas por sindicatos ou escritórios de advocacia de massa. Além dos impactos sociais e da perda patrimonial da vítima, o golpe pode ocasionar sérios reflexos institucionais como aumento de consultas, reclamações, tentativas de

responsabilização da administração pública e a desconfiança sobre a atuação dos procuradores e do sistema judicial tornam o tema ainda mais relevante para os órgãos de representação jurídica do Estado.

Com o avanço da inteligência artificial, esses golpes têm se tornado cada vez mais sofisticados, empregando áudios gerados artificialmente, vídeos hiper-realistas e interações automatizadas, que dificultam a identificação da fraude. Diante desse cenário, é imperativo que as Procuradorias se envolvam ativamente na discussão e implementação de medidas tecnológicas avançadas para combater essas práticas.

Este artigo propõe discutir respostas concretas ao problema, com foco na utilização de ferramentas de machine learning voltadas à detecção e prevenção de fraudes jurídicas digitais. A partir de modelos já desenvolvidos na literatura acadêmica internacional, busca-se adaptar soluções eficazes ao contexto brasileiro e, em especial, ao modus operandi do falso advogado.

O primeiro capítulo examina a estrutura do golpe, seus mecanismos de operação e o perfil das vítimas. O segundo capítulo trata da responsabilização criminal e civil dos envolvidos, bem como da regulação das plataformas de mensagens. Por fim, o terceiro capítulo apresenta propostas tecnológicas baseadas em machine learning que podem ser incorporadas pelas instituições públicas como instrumentos estratégicos de prevenção e resposta.

2. O GOLPE DOS FALSOS ADVOGADOS

O golpe dos falsos advogados constitui uma prática fraudulenta sofisticada, na qual indivíduos se passam por profissionais do Direito com o objetivo de extorquir vítimas, geralmente em contextos jurídicos de vulnerabilidade. Conforme divulgado pelo Tribunal Regional Federal da 3ª Região, esse esquema tem como alvo recorrente pessoas com processos judiciais em curso, notadamente aquelas que aguardam o pagamento de valores oriundos de decisões judiciais, oferecendo serviços ilegítimos com promessas de resolução célere ou de vantagens processuais mediante pagamento indevido (TRF3, 2025).

A tática dos golpistas envolve o contato direto com as vítimas — por telefone, e-mail ou redes sociais — apresentando-se como advogados ou representantes de escritórios fictícios, frequentemente valendo-se de documentos, áudios e imagens gerados por inteligência artificial. Para conferir credibilidade à abordagem, os criminosos utilizam dados reais extraídos de processos públicos ou, em muitos casos, nomes de advogados devidamente registrados na OAB (OAB, 2025).

Mediante falsas promessas de liberação de valores ou êxito na ação judicial, exigem pagamentos prévios a título de “custas”, “honorários”, “taxas administrativas” ou até mesmo “gratificações” indevidas a servidores ou magistrados. Documentos falsificados, incluindo intimações judiciais, são comumente empregados para legitimar a fraude.

Em 22 de maio de 2025, a Ordem dos Advogados do Brasil registrou 2.181 ocorrências formais relacionadas a esse tipo de golpe, segundo dados do seu canal de denúncias institucional. Todavia, há fortes indícios de subnotificação, uma vez que muitas vítimas se sentem constrangidas em relatar que realizaram pagamentos indevidos ou confiaram em interlocutores fraudulentos, o que reforça a necessidade de políticas públicas e ações institucionais de acolhimento, orientação e prevenção (TRF3, 2025).

A crescente sofisticação dos meios utilizados para a fraude, como deepfakes e documentos produzidos com IA generativa (como as Generative Adversarial Networks – GANs), torna a identificação da falsidade ainda mais difícil. Tais tecnologias possibilitam a clonagem de vozes humanas e a criação de imagens e vídeos altamente verossímeis, sendo empregadas para simular a identidade de advogados. Já foram registradas situações em que criminosos utilizaram vozes clonadas para entrar em contato com partes, simulando tratativas oficiais, com tom técnico e institucional, o que agrava o risco de comprometimento da imagem das instituições.

Conforme noticiado pela Forbes Brasil, a inteligência artificial vem sendo usada para gerar documentos jurídicos com aparência legítima, redigidos com linguagem técnico-jurídica convincente, dificultando o reconhecimento da fraude até mesmo por sistemas automatizados (KAUFLIN; MASON, 2023). Essas ferramentas permitem a produção em larga escala de petições, notificações e boletos falsos, incluindo elementos visuais como assinaturas, selos e carimbos digitais.

Um elemento comum a essas práticas é o phishing jurídico que se baseia na personalização das mensagens com dados reais dos processos ou das partes envolvidas. Esse método explora a confiança depositada em instituições legais e nos profissionais em nome do qual os criminosos falam, tornando a fraude mais convincente (MOREIRA, 2023). A sofisticação da fraude reforça a urgência de mecanismos de proteção institucional, como alertas oficiais, selos de verificação da identidade funcional e campanhas educativas sobre canais legítimos de contato com as Procuradorias.

Em síntese, embora o objetivo principal dos golpistas seja a extorsão de recursos, a crescente complexidade e disseminação dessas fraudes coloca em risco não apenas os jurisdicionados, mas também a própria legitimidade da atuação pública. As procuradorias estaduais, que exercem

papel essencial na interlocução entre o Estado e os cidadãos, assumem protagonismo na formulação de estratégias de prevenção, resposta institucional e conscientização social frente a essa nova modalidade de criminalidade digital.

3. AUSÊNCIA DE REGULAMENTAÇÃO ESPECÍFICA E DESAFIOS JURÍDICOS NA REPRESSÃO AO GOLPE DO FALSO ADVOGADO

Ainda não há, no ordenamento jurídico brasileiro, uma regulamentação específica que trate da repressão a golpes semelhantes aos do falso advogado, sobretudo diante do uso de tecnologias modernas, como inteligência artificial e engenharia social. A resposta jurídica a esse fenômeno, por ora, se apoia em dispositivos genéricos do ordenamento civil e penal, o que torna sua prevenção e punição desafiadoras.

No âmbito civil, as vítimas podem pleitear a reparação por danos morais e materiais com fundamento no artigo 186 do Código Civil, que impõe o dever de indenizar àquele que, por ação ou omissão voluntária, negligência ou imprudência, causar dano a outrem. No entanto, a efetividade dessa responsabilização é dificultada pela natureza anônima e descentralizada da fraude. Os criminosos, muitas vezes, operam utilizando identidades falsas e contas bancárias em nome de terceiros (as chamadas “contas laranjas”), o que compromete a rastreabilidade das operações e o sucesso da responsabilização civil.

Na esfera penal, os dispositivos mais frequentemente utilizados para enquadrar a conduta são os artigos 171 (estelionato), 299 (falsidade ideológica) e 168 (apropriação indébita) do Código Penal. Esses dispositivos possibilitam certa resposta repressiva, especialmente quando há falsificação de documentos ou obtenção fraudulenta de valores mediante engano da vítima. Contudo, a ausência de uma tipificação penal específica para fraudes com uso de inteligência artificial, bem como a complexidade dos meios empregados pelos criminosos, impõe entraves à investigação e à persecução penal, que muitas vezes esbarram em limitações técnicas e na falta de cooperação entre órgãos.

No que diz respeito à regulamentação das tecnologias utilizadas nos golpes, especialmente a inteligência artificial, ainda não há um marco legal consolidado. Tramita atualmente na Câmara dos Deputados o Projeto de Lei nº 2.338/2023, que institui o marco legal da inteligência artificial no Brasil, inspirado no modelo europeu. Embora a proposta avance no sentido de estabelecer princípios para o desenvolvimento e uso ético da IA, especialistas apontam a necessidade de aprimoramentos para equilibrar inovação e responsabilidade, de modo a prever mecanismos de

prevenção e repressão ao uso indevido dessa tecnologia. O debate legislativo, no entanto, ainda está em curso, sem previsão definida para votação (DISTRITO FEDERAL, 2025).

Do ponto de vista preventivo, as iniciativas existentes ainda são incipientes. As ações concentram-se, sobretudo, em campanhas de conscientização promovidas por instituições como a OAB e tribunais, orientando a população sobre como reconhecer tentativas de fraude e verificar a autenticidade da inscrição profissional de advogados por meio da plataforma ConfirmADV. Algumas entidades também têm incentivado o uso de múltiplos fatores de autenticação (MFA), com o objetivo de reforçar a segurança no acesso a sistemas sensíveis. Ainda assim, diante da crescente sofisticação dos golpes — com uso de inteligência artificial para falsificação de vozes, documentos e vídeos — tais medidas se revelam insuficientes para garantir uma proteção efetiva, especialmente diante do perfil das vítimas, muitas vezes vulneráveis ou em situação de fragilidade emocional.

Outro fator que contribui para a perpetuação do golpe é a fragmentação institucional. A ausência de articulação entre OAB, Ministério Público, forças policiais, Poder Judiciário e entidades de fiscalização dificulta a construção de respostas coordenadas e integradas, tanto no plano repressivo quanto no educativo. A falta de políticas públicas voltadas à educação digital da população e ao fortalecimento de canais seguros de comunicação institucional amplia o campo de atuação dos golpistas.

Em síntese, a repressão ao golpe do falso advogado ainda depende de interpretações analógicas de normas civis e penais, sem que haja, até o momento, uma estrutura normativa ou tecnológica adequada para lidar com fraudes estruturadas sobre bases tecnológicas avançadas. A atuação coordenada das instituições e a valorização de políticas preventivas, com papel destacado das procuradorias estaduais na formulação e disseminação de práticas seguras de interlocução com os jurisdicionados, são passos essenciais para conter os impactos dessa nova forma de criminalidade digital.

4. A INTELIGÊNCIA ARTIFICIAL COMO INSTRUMENTO DE COMBATE AO GOLPE DO FALSO ADVOGADO

Como visto nos tópicos anteriores, a crescente sofisticação dos crimes digitais tem imposto desafios significativos às instituições públicas e privadas, sobretudo em razão do uso disseminado e crescente de aplicativos de mensagens — em especial o WhatsApp, da empresa Meta, atualmente o mais utilizado no mundo:

> With the rise of Over-The-Top (OTT) platforms like Whatsapp, Telegram, etc., phishing messages have expanded their reach on these platforms. According to the statistics [4], Kaspersky Internet Security for Android detected that the most significant share of detected malicious links between December 2020 and May 2021 was sent via WhatsApp (89.6%), followed by Telegram (5.6%). Smishing attacks frequently request that the victim open a link, call a number, or send an email address the attacker has provided through an SMS message (HARICHANDANA et al., 2024).

No Brasil, entre os crimes que vêm se intensificando, destacam-se os golpes perpetrados por falsos advogados por meio de aplicativos de mensagens. Esse fenômeno apresenta complexidade elevada, sobretudo em razão do perfil das vítimas – geralmente pessoas em situação de vulnerabilidade – e da recente utilização de tecnologias baseadas em inteligência artificial (IA), capazes de gerar áudios e vídeos com elevado grau de verossimilhança.

Diante da limitada eficácia das estratégias tradicionais de prevenção, impõe-se a necessidade de desenvolver soluções tecnológicas que não apenas coíbam, mas desestimulem de forma proativa a prática dessas fraudes. Uma das possibilidades mais promissoras reside na aplicação de modelos de machine learning para a detecção de padrões linguísticos e comportamentais associados a fraudes jurídicas. Já existem, inclusive, modelos empregados com êxito em setores como o bancário (LIMA et al, 2024) e o comércio eletrônico (KANG, 2019), baseados em leitura semântica, análise de conteúdo e clusterização de mensagens suspeitas.

A literatura recente destaca propostas que combinam ontologias semânticas, conjuntos predefinidos de dados e algoritmos de IA para encriptação, decriptação e criptoanálise de comunicações suspeitas. Um exemplo relevante é o framework desenvolvido por HUSSAIN e MOHIDEEN (2024) que emprega sistemas criptográficos integrados a modelos preditivos capazes de identificar potenciais crimes por meio da análise de terminologias suspeitas em ambientes digitais.

> The outlined framework addresses this gap by employing a cryptographic system that integrates semantic web ontology, a Predefined Dataset, and machine learning technology for encryption, decryption, and cryptanalysis. A key feature is the framework's ability to predict potential crimes by analyzing blogs for suspicious terminology conveyed through coded information. The identified information is then promptly communicated to law enforcement agencies, thereby alleviating the burden on numerous security entities and offering an efficient strategy against global crime (HUSSAIN e MOHIDEEN, 2024)

No caso analisado, o sistema de machine learning atua de forma antecipatória, por meio da descriptografia e análise semântica de mensagens veiculadas em blogs e redes sociais, identificando termos suspeitos associados à prática de golpes. Uma vez detectadas essas mensagens, o sistema emite alertas automáticos às autoridades policiais, que devem, por sua vez, adotar medidas preventivas antes da consumação do delito. Trata-se, portanto, de um modelo preditivo, baseado em

inteligência artificial, que visa evitar a prática criminosa por meio da detecção precoce de padrões comportamentais e discursivos. No entanto, sua efetividade depende diretamente da existência de uma estrutura policial bem organizada, com protocolos claros e capacidade de resposta imediata às ameaças sinalizadas.

O artigo de TEJA et al., (2024) apresenta outro exemplo relevante, voltado à análise do comportamento de usuários em grupos de mensagens, com o objetivo de identificar padrões suspeitos de interação. Nesse caso, o foco recai sobre a dinâmica das conversas em ambientes coletivos, permitindo a antecipação de estratégias fraudulentas por meio da observação de recorrências linguísticas, estrutura de rede e intensidade de comunicação entre membros.

> Spam, which is generally regarded as junk or unsolicited information or messages, could come in the form of emails or text messaging. When spam attacks through one of these channels, there can be risks that involve leaks of personal information, invasions of privacy, or access to unauthorized data from mobile devices, meaning it is just as dangerous regardless.
> (…)
> The objective of this web application is to provide a comprehensive analysis of WhatsApp chats, including the total number of messages, a graph of monthly timelines, most frequently used words, and the active participants in the conversation.
> Additionally, the application aims to detect spam messages sent by individuals and display the total number of spam messages present in the chat. The proposed system is a web application that can analyze WhatsApp chats and provide useful insights. The system is implemented using simple Python modules such as Pandas, Steamline, Seaborn, and WordCloud, which are used to create data frames and plot different graphs. The system can be applied to the largest dataset efficiently, making it suitable for analyzing chats from both individuals and groups.

A proposta dos autores consiste em utilizar a ferramenta WhatsApp Chat Analyzer para identificar mensagens de spam enviadas em grupos, por meio da análise da atividade de chat. A partir dessa análise, é possível extrair informações relevantes, como os usuários mais ativos, os dias ou meses com maior volume de mensagens, e a linha do tempo diária de um determinado participante. Os dados fornecidos pelo próprio WhatsApp foram utilizados como ponto de partida para uma investigação mais aprofundada, com o intuito de identificar padrões de comportamento e o grau de participação dos indivíduos envolvidos. Os resultados permitiram visualizar a interação entre diferentes autores ao longo do tempo, bem como os emojis mais utilizados e outras métricas significativas de engajamento (TEJA et al., 2024).

Outra proposta voltada à detecção automática de mensagens de spam e phishing[1] foi apresentada por MANURUNG; MUNAWIR; PRADEKA, 2025, por meio do desenvolvimento de uma aplicação que utiliza os métodos TF-IDF e Random Forest para a classificação de mensagens no WhatsApp. A ferramenta foi implementada como um aplicativo Android, com o objetivo de filtrar comunicações suspeitas antes que elas cheguem ao destinatário, emitindo alertas ao usuário sobre possíveis riscos de fraude. A acurácia reportada superou 90%, evidenciando o potencial da aplicação de técnicas de machine learning na prevenção de crimes digitais em plataformas de mensagens instantâneas (MANURUNG; MUNAWIR; PRADEKA, 2025).

> The application employs the TF-IDF (Term Frequency-Inverse Document Frequency) method and machine learning using the Random Forest model. It is developed using the MVVM (Model-View ViewModel) architecture, enabling the separation of business logic from the user interface, thereby improving development and maintenance efficiency. The research findings demonstrate that the combination of TF-IDF and Random Forest achieves high accuracy in classifying spam and phishing messages.

Uma proposta mais direcionada à prevenção do golpe praticado por falsos advogados consistiria na implementação de um bloqueio automático de mensagens, impedindo que elas sequer fossem recebidas pela vítima. Tal medida teria o potencial de interromper a prática do estelionato ainda em sua fase preparatória, agindo de forma preventiva e eficaz. Considerando que a OAB mantém um Cadastro Nacional de Advogados[2], contendo dados como nome e número de telefone, é tecnicamente viável estabelecer uma integração entre esse banco de dados e os aplicativos de mensagens instantâneas, como o WhatsApp.

A partir dessa integração, seria possível desenvolver sistemas de machine learning capazes de identificar, em tempo real, números não cadastrados como pertencentes a advogados habilitados. Ao detectar a tentativa de contato a partir de um número não reconhecido, o sistema poderia bloquear automaticamente a mensagem antes de sua entrega ao destinatário, prevenindo o engano. Tal abordagem se alinha à lógica de um sistema antecipatório e automatizado, voltado à mitigação de riscos em ambientes digitais, especialmente no que se refere à proteção de pessoas mais vulneráveis a fraudes.

Contudo, tanto os sistemas descritos pelos autores acima quanto a proposta de bloqueio automático de mensagens enfrentam obstáculos significativos, que, nas condições atuais, se revelam

[1] Phishing, according to the APWG, involves stealing confidential and important information from individuals using advanced methods, techniques, and tools. Phishing can be executed via email (phishing email), online social media platforms, and mobile applications (MANURUNG et al., 2024).
[2] https://cna.oab.org.br/

praticamente insuperáveis do ponto de vista técnico e jurídico. O principal entrave reside no fato de que a leitura automática do conteúdo das mensagens, etapa necessária para a identificação de padrões fraudulentos, viola a política de criptografia de ponta a ponta adotada pelo WhatsApp, aplicativo de mensagens da empresa Meta, atualmente o mais utilizado para a prática desse tipo de golpe (NOIA, 2023).

Essa criptografia garante que somente o remetente e o destinatário tenham acesso ao conteúdo das mensagens, impedindo, em tese, qualquer leitura ou interceptação por terceiros, inclusive por sistemas automatizados. Embora já existam ferramentas tecnológicas capazes de quebrar ou decifrar textos criptografados com elevado grau de precisão — conforme demonstrado por HUSSAIN e MOHIDEEN (2024) —, sua utilização implicaria graves riscos à privacidade dos usuários e levantaria sérias implicações legais e éticas.

Outros problemas poderiam surgir como o bloqueio de advogados com cadastro recente ainda não sincronizados e ainda de servidores de cartórios, secretarias de justiça, estagiários e outros atores legítimos.

Dessa forma, a viabilidade de um sistema preventivo com bloqueio automático, tal como o proposto, depende não apenas de avanços tecnológicos, mas sobretudo de revisões regulatórias e consensos institucionais sobre os limites entre segurança digital e proteção de dados pessoais.

Contudo, apesar da atual inviabilidade técnica e regulamentar da implementação de um sistema de bloqueio automático de mensagens suspeitas enviadas por falsos advogados — sobretudo diante das limitações impostas pela criptografia de ponta a ponta e pela legislação de proteção de dados —, existem alternativas tecnológicas viáveis que podem contribuir para a prevenção de golpes dessa natureza. Uma dessas possibilidades é a criação de um sistema automatizado de alerta, que, ao receber uma mensagem de um número não vinculado a advogado regularmente inscrito na OAB, seja capaz de analisar o conteúdo textual e emitir advertência ao destinatário sobre o possível risco, sem bloquear ou censurar o contato.

Essa estratégia preserva tanto a integridade da comunicação quanto o respeito à privacidade e à liberdade de expressão, ao mesmo tempo em que informa preventivamente o usuário sobre a autenticidade do remetente.

HARICHANDANA et al. (2024), propôs uma arquitetura de sistema baseada exatamente nesse modelo, integrando ontologias da web semântica, técnicas de machine learning e um banco de dados verificado de profissionais. A ferramenta realiza a leitura e interpretação semântica das mensagens recebidas, correlacionando o conteúdo com padrões previamente identificados como

suspeitos, e gera alertas visuais ou sonoros ao usuário, sem interferir na entrega da mensagem. Além disso, o sistema possibilita o encaminhamento automatizado de notificações às autoridades competentes, caso o padrão identifique risco concreto de tentativa de estelionato.

> Figure 1 overviews the existing system and a system with COPS implemented. Many smishing messages come with a known brand name or associated icons. However, instead of the original website link, a phishing link is provided. Aware of the brand icon, the user clicks the link and gets phished. Instead, if a COPS is deployed, the user can be alerted in real-time about the potentially dangerous nature of the message, and he/she can make an informed decision.

Na Figura abaixo, os autores HARICHANDANA et al. (2024) apresentam a interface gráfica da aplicação em funcionamento, evidenciando o layout da ferramenta utilizada para a detecção em tempo real de mensagens suspeitas, por meio de processamento local e análise automatizada do conteúdo textual.

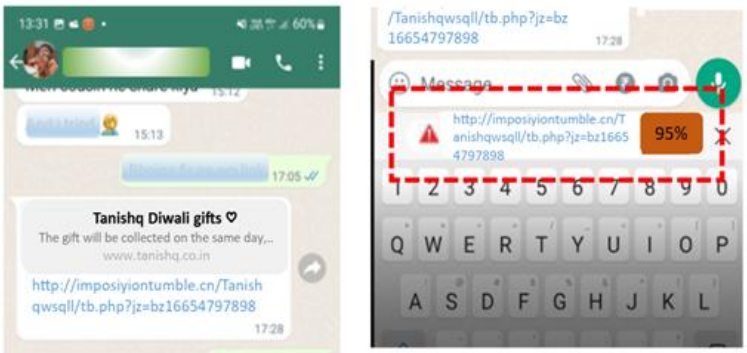

Fig. 1: Smishing Detection: To the left is the existing system, and to the right is a system with the COPS pipeline

A arquitetura do sistema proposta por HARICHANDANA et al. (2024) é composta por uma sequência de etapas integradas, concebidas para viabilizar a detecção antecipada de tentativas de fraude por meio de mensagens digitais:

> • Screen Understanding framework: Screen understanding framework has three major functionalities: – Understanding which application is currently displayed, based on the understanding, process just the screen related to the communication/messaging category. – Understand the views or user interface (UI) elements associated with the displayed screen. – Extract the texts related to views
> • Smishing Detection network: Verify the textual content provided by the Screen understanding framework. The model architecture is explained in detail in Section V.
> • Smishing Notifier: Once Smishing is detected with high probability, the smishing notifier can notify the user by providing notifications or inline suggestions.

No caso específico deste trabalho, o modelo de machine learning proposto funcionaria por meio do cruzamento de dados públicos disponíveis no Cadastro Nacional da OAB com notificações exibidas pelo sistema operacional do dispositivo móvel. Essa abordagem permite a identificação preliminar de mensagens potencialmente suspeitas, com base na origem do número do contato, sem violar a criptografia de ponta a ponta adotada pelo WhatsApp.

Dessa forma, o sistema não acessa diretamente o conteúdo completo das mensagens, mas atua apenas sobre os metadados e os fragmentos informativos que o próprio sistema operacional expõe ao usuário. Trata-se, portanto, de uma solução que busca compatibilizar a eficácia da prevenção de fraudes com o respeito à privacidade e à integridade das comunicações protegidas por criptografia.

Ao detectar uma nova interação no WhatsApp, o sistema verificaria se o número de origem está cadastrado como advogado ativo no banco de dados da OAB. Em caso negativo, e se identificada alguma linguagem de risco, o aplicativo exibiria um alerta como: “Atenção: este número não está vinculado a advogado registrado na OAB. Cuidado com possíveis tentativas de fraude.”

É importante destacar que esse sistema não exige o rompimento da criptografia do WhatsApp, uma vez que a análise é realizada a partir das notificações do sistema, de forma similar ao que já ocorre com aplicativos de segurança. A base de dados necessária já existe na OAB, sendo suficiente sua atualização periódica para incluir números de telefone celular vinculados a advogados ativos. O foco do sistema é a conscientização e a prevenção, e não a censura, o que reforça sua aderência aos princípios da liberdade de comunicação e do devido processo legal.

Complementarmente, sugere-se a criação de um selo de verificação profissional, denominado “OAB Verificada”, a ser incorporado ao WhatsApp Business. Assim como ocorre com perfis de empresas, os advogados e sociedades de advogados poderiam obter um selo oficial de identificação, validado pela OAB, reforçando a credibilidade do interlocutor.

Os benefícios desse modelo são diversos. Em primeiro lugar, fortalece-se a segurança jurídica, reduzindo de forma expressiva o risco de fraudes por falsos advogados. Em segundo, valoriza-se a identidade profissional da advocacia brasileira, conferindo maior legitimidade e confiabilidade ao exercício da profissão. Em terceiro, a proposta se mostra tecnologicamente viável, pois se apoia na infraestrutura já existente do WhatsApp Business e pode ser implementada gradualmente pelas seccionais da OAB em todo o território nacional.

Em síntese, a utilização de inteligência artificial no combate a fraudes praticadas por falsos advogados não apenas é possível, como se revela urgente e necessária. A proposta aqui

delineada oferece uma resposta proporcional ao desafio enfrentado, conjugando inovação tecnológica, segurança jurídica e respeito às garantias fundamentais. Para sua concretização, será imprescindível o engajamento articulado entre instituições públicas, setor privado e sociedade civil, sob o compromisso comum de preservar a integridade das relações jurídicas em ambiente digital.

5. CONCLUSÃO

O avanço das fraudes digitais exige a reformulação das estratégias de prevenção e resposta por parte das instituições públicas, em especial das Procuradorias dos Estados. A atuação das procuradorias estaduais revela-se essencial na construção de canais de comunicação confiáveis, na orientação preventiva aos jurisdicionados e na articulação com outras instituições para mitigar os impactos das fraudes.

A criação de alertas automáticos, a integração com bases de dados da OAB e o uso de selos de verificação profissional são exemplos de medidas tecnicamente viáveis e juridicamente compatíveis com a proteção de dados e a liberdade de comunicação.

Embora haja limitações decorrentes da criptografia ponta a ponta em aplicativos como o WhatsApp, é possível desenvolver soluções de alerta e orientação sem violar garantias fundamentais.

O uso estratégico da inteligência artificial, aliado à expertise da advocacia pública, pode representar um divisor de águas na proteção da população contra fraudes sofisticadas. A consolidação dessas medidas depende, porém, do engajamento político-institucional e do investimento em educação digital, infraestrutura tecnológica e regulamentação específica.

6. BIBLIOGRAFIA

BRASIL. Ministério da Fazenda. **Regulação equilibrada da inteligência artificial ajudará a fortalecer o novo ciclo de desenvolvimento do país**. Governo do Brasil, 2025. Disponível em: https://www.gov.br/fazenda/pt-br/assuntos/noticias/2025/junho/regulacao-equilibrada-da-inteligencia-artificial-ajudara-a-fortalecer-o-novo-ciclo-de-desenvolvimento-do-pais. Acesso em: 28 jun. 2025.

CAIXETA, Nayara; MACHADO, Rita. **Deepfake e crimes digitais: quando a realidade se torna uma arma**. Correio Braziliense, Brasília, maio 2025. Disponível em:

https://www.correiobraziliense.com.br/direito-e-justica/2025/05/7136758-deepfake-e-crimes-digitais-quando-a-realidade-se-torna-uma-arma.html. Acesso em: 28 jun. 2025.

DISTRITO FEDERAL. Agência Brasília. **Debate sobre regulação da IA movimenta a Campus Party**. Governo do Distrito Federal, 2025. Disponível em: https://www.agenciabrasilia.df.gov.br/w/debate-sobre-regula%C3%A7%C3%A3o-da-ia-movimenta-a-campus-party?redirect=%2Fnoticias%2F. Acesso em: 28 jun. 2025.

GOLDENBERG, Mirian. **No dia 1º de abril, quase caí no golpe do falso advogado**. Folha de S.Paulo, São Paulo, 1 abr. 2025. Disponível em: https://www1.folha.uol.com.br/colunas/miriangoldenberg/2025/04/no-dia-1o-de-abril-quase-cai-no-golpe-do-falso-advogado.shtml. Acesso em: 28 jun. 2025.

HARICHANDANA, B. S. S. et al. **COPS: a compact on-device pipeline for real-time smishing detection. 2024.** arXiv:2402.04173. Disponível em: https://arxiv.org/abs/2402.04173 . Acesso em: 27 jun. 2025.

HUSSAIN, Syed; MOHIDEEN, Pakkir. **A state-of-the-art universal machine learning framework for decoding suspect coded messages**. Measurement: Sensors, v. 33, p. 101115, 2024. Disponível em https://ui.adsabs.harvard.edu/abs/2024MeasS..3301115H/abstract. Acesso em 24 jun. 2025.

KANG, Haimeng. **Fraud detection in mobile money transactions using machine learning.** 2019. Thesis (Master of Science) – Iowa State University, Ames, 2019. Disponível em: https://dr.lib.iastate.edu/bitstreams/02d3b629-4b8a-4e2e-a498-7e0910636d5a/download . Acesso em: 27 jun. 2025.

KAUFLIN, Jeff; MASON, Emily. **O ChatGPT a serviço do crime: como a IA facilita as fraudes financeiras.** Forbes Brasil, set. 2023. Disponível em: https://forbes.com.br/forbes-money/2023/09/o-chatgpt-a-servico-do-crime-como-a-ia-facilita-as-fraudes-financeiras/. Acesso em: 27 jun. 2025.

KRISHNA, A.; SHAIK, Subhani et al. **WhatsApp chat analysis and spam discovery using machine learning models**. In: INTERNATIONAL CONFERENCE ON SMART COMPUTING AND SYSTEMS FOR SUSTAINABLE DEVELOPMENT (ICSCSP 2024). Springer, 2025. p. 1–12. Disponível em: https://link.springer.com/chapter/10.1007/978-981-96-0924-6_1 Acesso em: 27 jun. 2025.

LIMA, Juliana S. et al. **Financial fraud detection through the application of machine learning techniques: a systematic literature review.** Humanities and Social Sciences Communications, London, v. 11, art. 232, 2024. Disponível em: https://www.nature.com/articles/s41599-024-03606-0 . Acesso em: 27 jun. 2025.

MAHER, Carmel. **Friend or Fraud? What is a Deepfake and How Does it Impact Fraud?** Mitek Systems, 2024. Disponível em: https://www.miteksystems.com/blog/friend-or-fraud-what-is-a-deepfake-and-how-does-it-impact-fraud. Acesso em: 27 jun. 2025.

MANURUNG, Ferdinand Aprillian et al. **Spam and phishing WhatsApp message filtering application using TF-IDF and random forest**. GISA - Global Information and Software Architecture Journal, [S. l.], 2024. Disponível em: https://tecnoscientifica.com/journal/gisa/article/download/551/280. Acesso em: 27 jun. 2025.

MOREIRA, Julia. **Mais de 550 vítimas caem no golpe do falso advogado no Rio só neste ano.** CNN Brasil, Rio de Janeiro, 29 jun. 2023. Disponível em: https://www.cnnbrasil.com.br/nacional/sudeste/rj/mais-de-550-vitimas-caem-no-golpe-do-falso-advogado-no-rio-so-neste-ano/. Acesso em: 27 jun. 2025.

MORGADO, Eduardo Martins. **Caso de cyber fraud por telefone no Brasil e a inteligência artificial: vítimas idosas, spoofing até a manipulação por engenharia social.** Publicações, 2023. Disponível em: https://www.editorapublicar.com.br/ojs/index.php/publicacoes/article/view/794. Acesso em: 24 jun. 2025.

NOIA, Julia. **WhatsApp é campeão de fraudes na internet, mostra pesquisa. Saiba como se proteger.** 2023. Disponível em: https://extra.globo.com/economia-e-financas/whatsappcampeao-de-fraudes-na-internet-mostra-pesquisa-saiba-como-se-proteger-25415122.html. Acesso em: 25 jun. 2025.

ORDEM DOS ADVOGADOS DO BRASIL - Seção São Paulo (OAB SP). **Cartilha do Golpe do Falso Advogado**. São Paulo: Comissão de Direitos e Prerrogativas e Comissão de Fiscalização da Atividade Profissional, 2025. Disponível em: https://www.oabsp.org.br/upload/1164693296.pdf. Acesso em: 28 jun. 2025.

REDDY, Mothe Vikas; ASHRITH, Arvapally; HARISH, Vunnam**. WhatsApp chat analysis and spam message detection**. Journal of Emerging Technologies and Innovative Research (JETIR), [S. l.], v. 10, n. 3, p. 456–462, mar. 2023. Disponível em: https://www.jetir.org/view?paper=JETIR2303560. Acesso em: 27 jun. 2025.

SEGUNDO Hugo de Brito Machado**. Direito e Inteligência Artificial: o que os algoritmos têm a ensinar sobre interpretação**, valores e justiça. 2. ed. Indaiatuba, SP: Editora Foco, 2024.

SOUZA, Nelson. **Deepfakes, provas digitais e fraudes no INSS: o colapso da confiança**. ConJur, 28 de maio de 2025. Disponível em: https://www.conjur.com.br/2025-mai-28/deepfakes-provas-digitais-e-fraudes-no-inss-colapso-da-confianca/. Acesso em: 28 jun. 2025.

TEJA, Dornala et al. **Chat analysis and spam detection of WhatsApp using machine learning**. ResearchGate, 2024. Disponível em: https://www.researchgate.net/publication/380959336_Chat_Analysis_and_Spam_Detection_of_WhatsApp_Using_Machine_Learning. Acesso em: 27 jun. 2025.

TRIBUNAL DE JUSTIÇA DO ESTADO DE SÃO PAULO. **Cuidado com golpes em falsos leilões, telefonemas, mensagens e sites**. São Paulo, 2025. Disponível em: https://www.tjsp.jus.br/Noticias/Noticia?codigoNoticia=105793. Acesso em: 27 jun. 2025.

BRASIL. Tribunal Regional Federal da 3ª Região. **Golpe do falso advogado**. São Paulo, 2025. Disponível em: https://www.trf3.jus.br/campanhas/2025/golpe-falso-advogado. Acesso em: 7 jul. 2025.