REVISTA TÓPICOS
ISSN 2965-6672 | Qualis A2

DETECÇÃO DE FRAUDES DIGITAIS: UMA ABORDAGEM SOBRE INTERAÇÃO HUMANO-COMPUTADOR APLICADA À SEGURANÇA PÚBLICA

DIGITAL FRAUD DETECTION: A HUMAN-COMPUTER INTERACTION APPROACH APPLIED TO PUBLIC SECURITY

Ciências Sociais Aplicadas, Ciências Exatas e da Terra • 07/04/2026

REGISTRO DOI: 10.70773/revistatopicos/775536664

Marcelo Meirelles Rezende[1]

RESUMO

O aumento dos delitos virtuais, principalmente os que envolvem manipulação cognitiva e engenharia social, gera desafios importantes para as entidades de segurança pública e os usuários de ambientes digitais. Apesar de o campo da Interação Humano-Computador (IHC) ter progredido em áreas como usabilidade, acessibilidade e inteligência artificial, ainda existe uma lacuna significativa na implementação desse conhecimento em situações críticas, como as fraudes online.

Neste artigo, é apresentado um modelo de interfaces intencionais focado na detecção de indícios de manipulação durante interações digitais, unindo conceitos de design centrado no usuário, visualização de dados e identificação automatizada de padrões. Em termos metodológicos, a pesquisa combina uma revisão teórica interdisciplinar, a elaboração de um modelo conceitual e a sugestão de um experimento com os usuários.

Como contribuição, este trabalho apresenta a identificação de intenção maliciosa como uma nova vertente analítica em IHC, expandindo o escopo para usos em segurança pública e na tomada de decisões em situações de risco.

Palavras-chave: Interação Humano-Computador; Segurança Pública; Fraudes Digitais; Inteligência Artificial; Desinformação.

ABSTRACT

The rise in cybercrimes, particularly those involving cognitive manipulation and social engineering, poses significant challenges for public safety agencies and users of digital environments. While the field of Human-Computer Interaction (HCI) has progressed in areas such as usability, accessibility, and artificial intelligence, a significant gap remains in the implementation of this knowledge in critical situations, such as online fraud.

This article presents a model of intentional interfaces focused on detecting signs of manipulation during digital interactions, combining concepts of user-centered design, data visualization, and automated pattern identification. Methodologically, the research combines an interdisciplinary theoretical review, the development of a conceptual model, and the suggestion of a user experiment.
As a contribution, this work presents the identification of malicious intent as a new analytical approach in HCI, expanding its scope for uses in public safety and decision-making in high-risk situations.
Keywords: Human-Computer Interaction; Public Safety; Digital Fraud; Artificial Intelligence; Disinformation.

1. INTRODUÇÃO

A atual transformação digital tem aumentado de maneira considerável a complexidade das relações sociais, econômicas e institucionais, trazendo tanto novas oportunidades quanto novas fragilidades. Dentre essas fragilidades, destacam-se os delitos cibernéticos que se baseiam em manipulação psicológica, como fraudes online e engenharia social, os quais aproveitam as vulnerabilidades cognitivas dos usuários.

O crescimento de sistemas algorítmicos, inteligência artificial e plataformas digitais tem intensificado a influência da mediação tecnológica nas interações sociais, impactando de maneira direta os processos decisórios. Recentes evidências empíricas revelam que a maneira como as informações são exibidas nas interfaces exerce um efeito significativo na percepção de riscos e nas escolhas dos usuários.

No domínio da Interação Humano-Computador, avanços importantes têm sido alcançados em áreas como usabilidade, acessibilidade e experiência do usuário. Contudo, investigações atuais sinalizam que a combinação de inteligência artificial e design focado no usuário ainda encontra limitações, especialmente em situações de alta criticidade. Informações recentes do setor financeiro elucidam a gravidade dos delitos digitais no Brasil. Uma pesquisa nacional revelou que cerca de 38% dos cidadãos brasileiros já foram afetados por fraudes ou tentativas de fraudes ligadas a serviços bancários, destacando a ampla prevalência desse fenômeno na sociedade atual.

Ademais, nota-se que a vulnerabilidade não é exclusiva de grupos particulares, atingindo diversas idades, níveis educacionais e perfis socioeconômicos. Este panorama sugere que a ocorrência de fraudes não pode ser atribuída apenas a fatores individuais, mas está interligada a elementos estruturais das interações digitais, como a forma em que as informações são apresentadas e compreendidas pelos usuários.

Nesse cenário, é vital examinar estratégias que ajudem a mitigar essas vulnerabilidades, especialmente por meio do design de interfaces que favoreçam a identificação de indícios de manipulação e risco. A crescente ocorrência de fraudes digitais revela a dimensão estrutural desse fenômeno no Brasil. Dados recentes informam que, apenas em janeiro, mais de 375 mil tentativas de fraudes foram registradas no país, evidenciando a elevada frequência e a disseminação dessas práticas no ambiente digital.

Esse contexto, associado ao fato de que uma parte significativa da população já foi sujeita a golpes ou tentativas de fraudes, reforça a

perspectiva de que a vulnerabilidade dos usuários não pode ser considerada um fenômeno isolado ou individual, mas sim como resultado de dinâmicas sistêmicas que abrangem tecnologia, comportamento e a forma como as interações digitais são estruturadas.

Diante desta situação, se torna crucial o desenvolvimento de soluções que atuem diretamente durante a interação, ajudando os usuários a discernir sinais de risco e a interpretar intenções que possam ser maliciosas.

Neste contexto, este artigo tem como objetivo responder à seguinte pergunta de pesquisa: de que maneira as interfaces digitais podem ser projetadas para ajudar os usuários a reconhecer intenções maliciosas em interações online?

Este estudo acrescenta à literatura ao:

I. apresentar um modelo conceitual de interface intencional voltado para a detecção de fraudes digitais;

II. integrar conceitos de inteligência artificial e design centrado no usuário em cenários críticos; e

III. expandir o campo da Interação Humano-Computador ao incluir a detecção de intenção maliciosa como uma nova dimensão de análise.

2. FUNDAMENTAÇÃO TEÓRICA

2.1. IHC e Mediação Cognitiva

A Interação humano-computador envolve sistemas digitais que atuam como mediadores da vivência humana, afetando diretamente a forma como as informações são percebidas e compreendidas. Dessa forma, as interfaces moldam a interação de maneira cognitiva, influenciando decisões e comportamentos. A literatura clássica sobre Interação humano-computador ressalta que a interface é o principal ponto de contato entre o usuário e o sistema, sendo crucial para a qualidade da interação e o desempenho humano. Nesse aspecto, interfaces mal elaboradas podem provocar erros, aumentar a carga cognitiva e prejudicar a eficácia das decisões, podendo até resultar em consequências graves em certos contextos. Em contrapartida, sistemas altamente usáveis tendem a diminuir a frequência de erros e aumentar a produtividade dos usuários.

Esse entendimento reforça a importância de projetar interfaces que não apenas exibam informações, mas que também ajudem ativamente na interpretação de sinais e na prevenção de decisões errôneas, especialmente em situações como fraudes digitais, onde erros cognitivos podem ser maliciosamente manipulados.

2.2. Comunicação de Intenção em Interfaces

Novas abordagens mostram que interfaces podem ser concebidas para transmitir intenções, utilizando componentes visuais e semânticos que ajudam a eliminar ambiguidades e facilitar a compreensão das mensagens. Pesquisas recentes sobre delitos cibernéticos envolvendo crianças e adolescentes demonstram que práticas como o aliciamento online seguem padrões estruturais de interação, onde o agressor cria laços progressivamente confiáveis com a vítima, usando estratégias comunicativas adaptadas aos

interesses do usuário. Esse procedimento, classificado como engenharia social, ilustra que o sucesso das ações criminosas está intimamente ligado à habilidade de manipular a percepção e interpretação da vítima no ambiente digital. Além disso, observa-se que a maior exposição às tecnologias digitais, especialmente após a pandemia, ampliou consideravelmente as oportunidades para esse tipo de crime, evidenciando a inadequação de abordagens que se baseiam exclusivamente em legislação e rastreamento técnico. Nesse contexto, é essencial desenvolver mecanismos que atuem diretamente na interação, ajudando os usuários a reconhecer padrões de risco e intenções maliciosas durante o uso de sistemas digitais.

2.3. Inteligência Artificial, Confiança e Interpretabilidade

A literatura contemporânea sublinha que a confiança em sistemas de inteligência artificial está ligada à transparência, previsibilidade e à capacidade explicativa dos modelos. As interfaces têm um papel central nesse processo, ao converter resultados algorítmicos em representações que sejam compreensíveis.

2.4. Integração Entre IA e Design Centrado no Usuário

Pesquisas recentes em nível internacional mostram que sistemas que se baseiam em aprendizado de máquina ainda têm uma forte ênfase técnica, com pouca integração dos princípios de design centrado no usuário. Essa defasagem impacta negativamente a interpretabilidade e a usabilidade dos sistemas.

2.5. Tomada de Decisão e Influência dos Sistemas Digitais

Estudos empíricos mostram que os sistemas digitais têm um impacto direto nas decisões tomadas, especialmente em contextos onde há assimetria de informação ou manipulação de conteúdo. Isso reforça a necessidade de interfaces que ofereçam suporte à interpretação crítica dos usuários.

2.6. Lacuna em Contextos Críticos

Apesar dos progressos, há uma notável falta de investigações voltadas à segurança pública e fraudes digitais, ressaltando a necessidade de abordagens que combinem detecção automática e assistência à decisão humana.

3. METODOLOGIA

A investigação utiliza uma abordagem qualitativa e exploratória, incluindo a proposta de um modelo conceitual e um desenho experimental.

3.1 Etapas

- Exame de conteúdo de mensagens enganosas
- Modelagem de comportamentos de intenção maliciosa
- Criação de um protótipo conceitual
- Sugestão de experimento comparativo com participantes

3.2 Métricas

- Taxa de detecção de fraudes

- Duração da decisão
- Percepção de ameaça

3.3. Validação Proposta

O desenho experimental é pautado em metodologias já estabelecidas em Interação Humano-Computador, possibilitando validações futuras por meio de experimentos controlados com usuários. A estrutura do modelo é desenhada para permitir comparações entre interfaces convencionais e interfaces aprimoradas.

4. MODELO PROPOSTO

O modelo consiste em três níveis:

- Nível analítico: análise de dados e identificação de padrões
- Nível visual: exibição de indicadores de risco e notificações
- Nível cognitivo: apoio à interpretação e à decisão

5. RESULTADOS ESPERADOS

- Aumento na detecção de fraudes
- Diminuir o tempo necessário para a decisão
- Reforço na percepção de risco

6. DISCUSSÃO

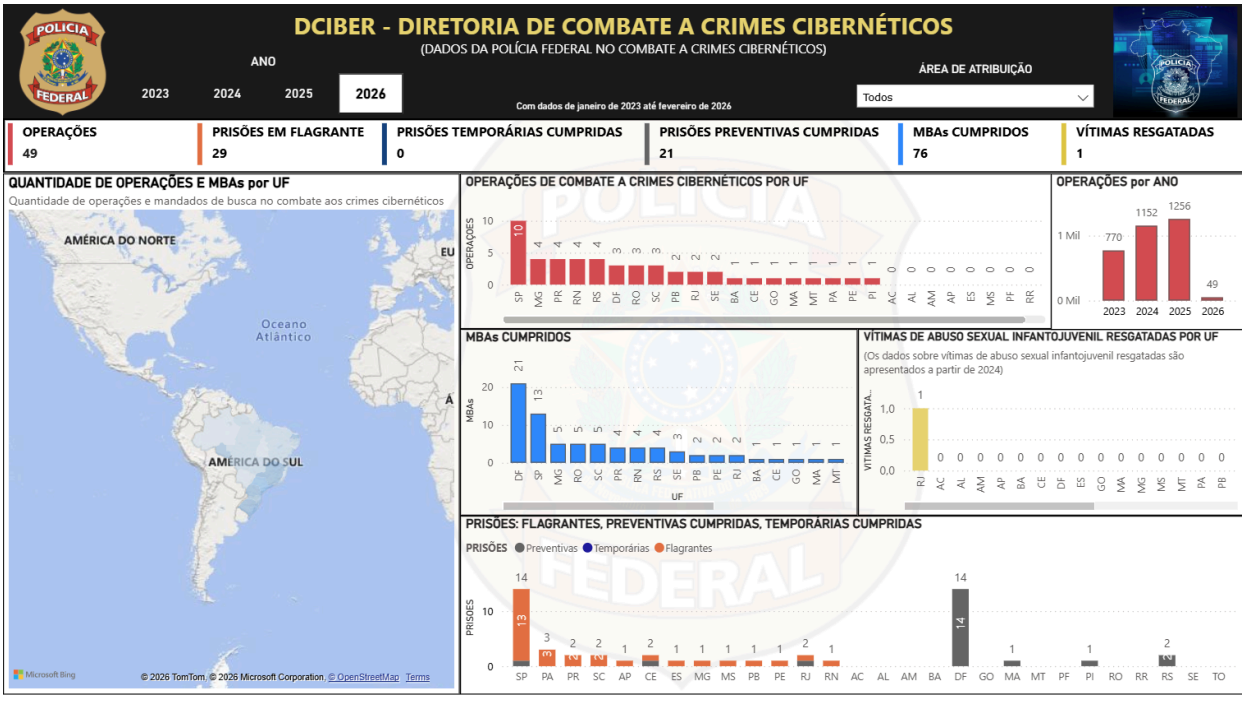
A proposta central deste trabalho defende que a identificação automática de padrões não é suficiente para auxiliar a tomada de decisão em contextos digitais complexos, sendo essencial o uso de interfaces que tornem as intenções nas interações visíveis e compreensíveis.

O modelo sugerido expande o domínio da Interação Humano-Computador ao integrar inteligência artificial e apoio cognitivo nas mais diversas situações críticas, favorecendo aplicações na segurança pública. Relatórios recentes do setor de segurança cibernética revelam que a maior parte das fraudes modernas não se fundamenta apenas em falhas tecnológicas, mas sim na manipulação de comportamentos humanos. Pesquisas de mercado indicam que criminosos utilizam táticas avançadas de engenharia social, explorando padrões cognitivos, confiança e reações.

Dados operacionais de sistemas de vigilância mostram que a incidência e a variedade das fraudes digitais estão aumentando, evidenciando padrões dinâmicos e a evolução constante das táticas criminosas. Painéis analíticos que se baseiam em informações institucionais indicam que os golpes digitais apresentam variações temporais e adaptam-se constantemente a novos contextos tecnológicos, complicando sua detecção através de métodos tradicionais.

Essas observações destacam que o combate às fraudes digitais requer métodos que transcendam a simples detecção

técnica, incorporando mecanismos que ajudem o usuário a interpretar sinais de risco enquanto interage com sistemas digitais. Segue analise da POLICIA FEDERAL de fraudes e demais crimes cibernéticos entre 2023-2025:



Fonte: https://www.gov.br/pf/pt-br/acesso-a-informacao/estatisticas/diretoria-de-combate-a-crimes-ciberneticos-dciber

É possível notar que métodos clássicos de identificação, que se baseiam somente em padrões técnicos, enfrentam dificuldades devido à crescente sofisticação dos criminosos no uso de ferramentas e plataformas digitais no cometimento de fraudes, levando a necessidade real de avanço de estratégias fundamentadas na análise do comportamento e confiabilidade das aplicações informáticas. Esse contexto ressalta a necessidade de soluções que combinem tecnologia com a compreensão do comportamento humano, especialmente por meio de interfaces que ajudem na detecção de sinais de risco durante as interações.

7. CONCLUSÃO

Este trabalho teve início com a percepção de que as fraudes digitais deixaram de ser meros eventos isolados para se tornarem fenômenos estruturais do atual ecossistema informacional. Dados empíricos provenientes de organizações como a Federação Brasileira de Bancos e a Serasa Experian evidenciam não somente a extensão dessas ocorrências, mas também sua repetição sistemática, sugerindo que o espaço digital está cada vez mais se estabelecendo como um ambiente de risco contínuo para a tomada de decisões.

No aspecto teórico, a ligação entre a Interação Humano-Computador, a inteligência artificial e as investigações sobre comportamento informacional revelou que a fragilidade dos usuários não resulta apenas de limitações pessoais, mas sim de um arranjo sociotécnico onde interfaces, algoritmos e estruturas comunicacionais funcionam como mediadores ativos da percepção e da ação. Dentro desse contexto, a literatura relacionada à confiança em sistemas automatizados sugere que a eficácia das tecnologias digitais está diretamente relacionada à sua habilidade de tornar seus processos decisórios compreensíveis, uma condição que ainda não é suficientemente atendida nos sistemas existentes.

Além disso, a inclusão de contribuições da criminologia e da Ciência da Informação demonstrou que atividades como aliciamento digital e cyberbullying ocorrem através de dinâmicas interativas em progressão, fundamentadas na construção de confiança, na diminuição de barreiras críticas e na exploração de viéses cognitivos. Essas evidências ressaltam que a fraude digital não é somente uma questão técnica, mas um fenômeno comunicacional e interpretativo, onde a manipulação da intenção desempenha um papel central.

A contribuição primordial deste artigo reside, assim, na proposta do conceito de interfaces intencionais como uma ampliação analítica e prática do campo da Interação Humano-Computador. Ao integrar a detecção de intenção maliciosa como uma dimensão essencial da interação, o modelo sugerido desloca a atenção da simples usabilidade para a mediação cognitiva em cenários de risco, incorporando processamento algorítmico, visualização de dados e suporte à decisão humana. Essa abordagem não só aprimora a eficácia dos sistemas digitais, mas também transforma o papel da interface em um componente ativo na redução de vulnerabilidades.

De uma perspectiva crítica, defende-se que abordagens que se concentram exclusivamente em detecção automatizada ou em respostas legais reativas são geralmente insuficientes diante da crescente sofisticação das táticas de engenharia social. A falta de mecanismos que auxiliem o usuário na interpretação dos sinais de risco perpetua um cenário onde a assimetria informacional favorece agentes maliciosos. Portanto, a incorporação de princípios de interpretabilidade e transparência nas interfaces surge como uma necessidade crucial para fortalecer a autonomia decisional dos usuários.

Finalmente, destaca-se que o combate às fraudes digitais requer uma abordagem integrada e interdisciplinar, que reúna tecnologia, design, educação e políticas públicas. Dentro desse cenário, o modelo sugerido apresenta um potencial para aplicação em ambientes institucionais, especialmente no campo da segurança pública, ao fornecer subsídios para desenvolver sistemas que não apenas identifiquem ameaças, mas também capacitem os usuários a reconhecê-las.

Como um plano para o futuro, sugere-se que sejam conduzidas pesquisas empíricas para confirmar o modelo em várias situações de aplicação, assim como a análise de sua conexão com sistemas autênticos de identificação de fraudes. Além disso, estudos sobre a efetividade de várias abordagens para a visualização e a comunicação de riscos podem ajudar a aprimorar a proposta, estabelecendo a detecção de intenções como um novo foco de investigação em Interação Humano-Computador em cenários críticos.

REFERÊNCIAS BIBLIOGRÁFICAS

ASSOCIATION FOR COMPUTING MACHINERY (ACM). User-centered design for machine learning systems. New York: ACM, 2024.

BARBOSA, Simone Diniz Junqueira; SILVA, Bruno Santana da. Interação humano-computador. Rio de Janeiro: Elsevier, 2010.

DE SOUZA, Clarisse Sieckenius. The semiotic engineering of human-computer interaction. Cambridge: MIT Press, 2005.

FEBRABAN. Relatório Radar Febraban 2025. São Paulo: Federação Brasileira de Bancos, 2025. Disponível em: https://cmsarquivos.febraban.org.br. Acesso em: 2026.

KAHNEMAN, Daniel. Thinking, fast and slow. New York: Farrar, Straus and Giroux, 2011.

NORMAN, Donald A. The design of everyday things. New York: Basic Books, 2013.

NATURE. Trust in AI: progress, challenges, and future directions. Humanities and Social Sciences Communications, 2024. Disponível em: https://www.nature.com. Acesso em: 2026.

REIS, Julio Cesar dos et al. Intenticons: communicating user intent through icons in interfaces. In: Proceedings of the ACM Conference, 2024.

RUSSELL, Stuart; NORVIG, Peter. Artificial intelligence: a modern approach. 4. ed. Pearson.

SPRINGER. Trust in artificial intelligence systems. Cham: Springer, 2025.

SCIENCEDIRECT. Human decision-making in digital environments. Journal of Behavioral Decision Making, 2024. Disponível em: https://www.sciencedirect.com. Acesso em: 2026.

SCIENCEDIRECT. Cognitive biases and digital interaction effects. Computers in Human Behavior Reports, 2024. Disponível em: https://www.sciencedirect.com. Acesso em: 2026.

SCIENCEDIRECT. Human-system interaction and decision behavior. Technology in Society, 2024. Disponível em: https://www.sciencedirect.com. Acesso em: 2026.

SCIENCE OF INFORMATION (SciELO). Mediação da informação e comportamento informacional. Perspectivas em Ciência da Informação, 2024. Disponível em: https://www.scielo.br. Acesso em: 2026.

SHNEIDERMAN, Ben et al. Designing the user interface: strategies for effective human-computer interaction. 6. ed. Boston: Pearson, 2016.

SERASA EXPERIAN. *Brasileiros sofreram mais de 375 mil tentativas de fraude em janeiro*. Disponível em: https://www.serasaexperian.com.br. Acesso em: 2026.

SPRINGER. Game development software engineering process life cycle: a systematic review. Journal of Computer Science, 2016. Disponível em: https://link.springer.com. Acesso em: 2026.

THALER, Richard H.; SUNSTEIN, Cass R. Nudge: improving decisions about health, wealth, and happiness. New Haven: Yale University Press, 2008.

---

[1] Mestrando em Sociologia Política Universidade Vila Velha (UVV).
ORCID: https://orcid.org/0009-0006-2336-874X