RCMOS – Revista Científica Multidisciplinar O Saber.
ISSN: 2675-9128. São Paulo-SP.

Ano VII, v.2 2026 | submissão: 03/07/2026 | aceito: 07/07/2026 | publicação: 17/08/2026

Deepfakes, Desinformação e Segurança Digital: Desafios para a Sociedade Conectada

Deepfakes, Disinformation and Digital Security: Challenges for the Connected Society

Alisson Garcia

Resumo

A rápida evolução da inteligência artificial generativa transformou significativamente a produção e a disseminação de conteúdos digitais. Entre essas inovações, destacam-se os deepfakes, definidos como mídias sintéticas produzidas por algoritmos capazes de reproduzir imagens, vídeos e áudios altamente realistas. Embora essa tecnologia apresente aplicações legítimas em áreas como entretenimento, educação e saúde, sua utilização maliciosa representa uma ameaça crescente à segurança digital, à privacidade, à democracia e à confiança social. Este estudo tem como objetivo analisar os impactos dos deepfakes na disseminação da desinformação e discutir os principais desafios à segurança digital em uma sociedade altamente conectada. Trata-se de uma pesquisa bibliográfica de abordagem qualitativa, baseada na literatura nacional e internacional. Os resultados demonstram que a combinação entre inteligência artificial generativa, redes sociais e velocidade de compartilhamento potencializa fraudes financeiras, manipulação política, ataques à reputação e crimes contra a identidade digital. Conclui-se que o enfrentamento desse fenômeno exige uma abordagem multidisciplinar que envolva inovação tecnológica, regulamentação jurídica, educação digital e cooperação internacional.

Palavras-chave: Deepfake; Inteligência Artificial; Desinformação; Segurança Digital; Cibersegurança.

Abstract

The rapid evolution of generative artificial intelligence has significantly transformed the production and dissemination of digital content. Among these innovations, deepfakes stand out, defined as synthetic media produced by algorithms capable of reproducing highly realistic images, videos, and audio. While this technology has legitimate applications in areas such as entertainment, education, and healthcare, its malicious use poses a growing threat to digital security, privacy, democracy, and social trust. This study analyzes the impact of deepfakes on the spread of disinformation and discusses the main challenges to digital security in a highly connected society. This is a bibliographic research with a qualitative approach, based on national and international literature. The results show that the combination of generative artificial intelligence, social networks, and the speed of sharing increases financial fraud, political manipulation, attacks on reputation, and crimes against digital identity. It is concluded that confronting this phenomenon requires a multidisciplinary approach involving technological innovation, legal regulation, digital education, and international cooperation.

Keywords: Deepfake; Artificial Intelligence; Disinformation; Digital Security; Cybersecurity.

1 Introdução

A transformação digital intensificou a produção e o compartilhamento de informações em escala global. Paralelamente aos benefícios proporcionados pelas tecnologias digitais, surgiram novas

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

RCMOS – Revista Científica Multidisciplinar O Saber.
ISSN: 2675-9128. São Paulo-SP.

Ano VII, v.2 2026 | submissão: 03/07/2026 | aceito: 07/07/2026 | publicação: 17/08/2026

modalidades de ameaças cibernéticas capazes de comprometer a confiança nas informações disponíveis na internet.

Entre essas ameaças, os deepfakes ocupam posição de destaque. Utilizando técnicas avançadas de aprendizado profundo (Deep Learning), especialmente Redes Generativas Adversariais (GANs) e modelos generativos mais recentes, esses sistemas conseguem produzir conteúdos audiovisuais extremamente convincentes, dificultando sua distinção em relação aos materiais autênticos.

A crescente sofisticação dessas ferramentas amplia significativamente o potencial de disseminação de notícias falsas, de manipulação da opinião pública, de fraudes financeiras e de ataques à reputação de indivíduos e instituições. O problema deixa de ser exclusivamente tecnológico e passa a representar um desafio social, jurídico, político e ético.

Nesse contexto, compreender os impactos dos deepfakes torna-se fundamental para fortalecer a segurança digital e a confiança nas informações compartilhadas em ambientes virtuais.

2 Deepfakes e a evolução da Inteligência Artificial

O termo deepfake resulta da combinação das palavras deep learning e fake, sendo utilizado para designar conteúdos sintéticos produzidos por inteligência artificial capazes de simular pessoas reais com elevado grau de realismo.

Inicialmente restritos a laboratórios de pesquisa, esses sistemas tornaram-se amplamente acessíveis com a popularização de softwares gratuitos e de modelos de código aberto. Atualmente, qualquer usuário com conhecimentos técnicos básicos pode gerar vídeos ou áudios falsificados em poucos minutos.

Os avanços recentes da Inteligência Artificial Generativa elevaram significativamente a qualidade dessas produções, reduzindo as imperfeições antes utilizadas para identificá-las. Consequentemente, tanto especialistas quanto usuários comuns enfrentam dificuldades crescentes em distinguir conteúdos verdadeiros de conteúdos manipulados. Embora existam aplicações legítimas na indústria cinematográfica, na acessibilidade, na preservação histórica e na educação, o uso indevido de deepfakes tem preocupado governos, pesquisadores e organismos internacionais.

3 Deepfakes como instrumento de desinformação

A desinformação caracteriza-se pela produção ou disseminação deliberada de conteúdos falsos com o objetivo de manipular percepções, comportamentos ou decisões coletivas. Os deepfakes ampliam significativamente esse fenômeno ao explorarem um dos principais mecanismos cognitivos humanos:

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

RCMOS – Revista Científica Multidisciplinar O Saber.
ISSN: 2675-9128. São Paulo-SP.

Ano VII, v.2 2026 | submissão: 03/07/2026 | aceito: 07/07/2026 | publicação: 17/08/2026

a confiança na evidência audiovisual. Historicamente, vídeos e áudios eram considerados provas relativamente confiáveis. Com a inteligência artificial, essa premissa passou a ser questionada.

No contexto político, vídeos manipulados podem influenciar processos eleitorais, estimular polarizações e enfraquecer instituições democráticas. Em ambientes corporativos, criminosos utilizam clonagem de voz para simular executivos e autorizar transferências financeiras fraudulentas. Já na esfera pessoal, observa-se um aumento de crimes envolvendo pornografia sintética, extorsão, chantagem e difamação.

Outro efeito preocupante é a chamada "erosão da confiança", fenômeno em que a população passa a desconfiar até mesmo de conteúdos autênticos, o que dificulta a construção de consensos sociais baseados em evidências.

4 Impactos na segurança digital

A segurança digital tradicionalmente concentrou-se na proteção de sistemas computacionais, redes e bases de dados. Entretanto, a ascensão dos deepfakes amplia esse conceito para incluir a proteção da identidade, da informação e da confiança digital.

Entre os principais impactos destacam-se:

- fraudes bancárias mediante clonagem de voz;
- ataques de engenharia social;
- roubo de identidade digital;
- manipulação de mercados financeiros;
- espionagem corporativa;
- comprometimento da reputação institucional;
- ataques à credibilidade da imprensa;
- disseminação acelerada de golpes virtuais.

Esses riscos são potencializados pela velocidade das redes sociais, nas quais conteúdos manipulados podem alcançar milhões de usuários antes mesmo de qualquer processo de verificação. A literatura também aponta que a evolução da IA torna os sistemas tradicionais de detecção cada vez menos eficientes, o que impulsiona uma corrida tecnológica entre algoritmos geradores e detectores.

5 Aspectos éticos e jurídicos

Os deepfakes suscitam importantes debates éticos relacionados à privacidade, ao consentimento, aos direitos da personalidade e à responsabilidade por conteúdos produzidos por inteligência artificial.

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

RCMOS – Revista Científica Multidisciplinar O Saber.
ISSN: 2675-9128. São Paulo-SP.
Ano VII, v.2 2026 | submissão: 03/07/2026 | aceito: 07/07/2026 | publicação: 17/08/2026

Do ponto de vista jurídico, diversos países discutem mecanismos regulatórios voltados à identificação obrigatória de conteúdos sintéticos, à responsabilização das plataformas digitais e à criminalização do uso malicioso dessas tecnologias.

No Brasil, embora não exista legislação específica dedicada exclusivamente aos deepfakes, diversas condutas podem ser enquadradas em dispositivos relacionados à honra, à fraude eletrônica, à proteção de dados pessoais e a crimes cibernéticos (Mirsky; Lee, 2020). Além disso, no âmbito dos processos eleitorais, a Justiça Eleitoral estabeleceu normas específicas para coibir o uso abusivo da inteligência artificial em campanhas políticas. A regulamentação, entretanto, deve equilibrar a proteção social, a inovação tecnológica e a liberdade de expressão, evitando restrições desproporcionais ao desenvolvimento científico.

6 Estratégias de enfrentamento

O combate aos impactos dos deepfakes exige ações integradas entre governos, empresas, pesquisadores e sociedade civil. Entre as principais estratégias destacam-se:

- desenvolvimento de algoritmos automáticos de detecção;
- certificação de autenticidade de conteúdos digitais;
- marcação de conteúdo produzido por IA;
- fortalecimento da alfabetização midiática;
- cooperação internacional entre plataformas digitais;
- atualização constante da legislação;
- campanhas permanentes de conscientização em segurança digital.

Além disso, a educação digital desempenha um papel central ao estimular o pensamento crítico e reduzir a vulnerabilidade da população diante de conteúdos manipulados.

Considerações finais

Os deepfakes representam um dos maiores desafios contemporâneos para a segurança da informação e a preservação da confiança nas relações digitais. A facilidade de produção, aliada à crescente sofisticação dos modelos de inteligência artificial, amplia significativamente o potencial de manipulação informacional em escala global.

Embora essa tecnologia apresente aplicações socialmente benéficas, seu uso indevido evidencia a necessidade de respostas coordenadas que envolvam inovação tecnológica, políticas públicas, regulamentação jurídica e educação digital. O enfrentamento desse cenário depende não apenas da

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.

Ano VII, v.2 2026 | submissão: 03/07/2026 | aceito: 07/07/2026 | publicação: 17/08/2026

evolução dos mecanismos de detecção, mas também da formação de cidadãos capazes de avaliar criticamente as informações consumidas no ambiente digital.

Dessa forma, fortalecer a alfabetização digital e promover uma cultura de responsabilidade no uso da inteligência artificial constituem medidas essenciais para preservar a integridade da informação e proteger a sociedade conectada diante das novas ameaças cibernéticas.

Referências

Mirsky Y, Lee W. The Creation and Detection of Deepfakes: A Survey. 2020.

Shoaib MR, Wang Z, Ahvanooey MT, Zhao J. Deepfakes, Misinformation, and Disinformation in the Era of Frontier AI. 2023.

Masood M, Nawaz M, Malik KM, Javed A, Irtaza A. Deepfake Generation and Detection. 2021.

Nguyen TTH, Nguyen QVH, Nguyen DT, et al. Deep Learning for Deepfakes Creation and Detection: A Survey. 2019.

Tavares CM. Inteligência Artificial e Deepfakes: Desafios Jurídicos e Tecnológicos à Integridade do Processo Democrático.

Este é um artigo publicado em acesso aberto (Open Access) sob a licença CreativeCommons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.