



DEEPPAKES E A PORNOGRAFIA ARTIFICIAL NÃO CONSENSUAL

Maria Luiza de Souza SANTOS¹

RESUMO: O presente artigo tem como objetivo discutir sobre as deepfakes e sua intrínseca relação com a pornografia artificial não consensual, a partir do contexto de expansão das tecnologias generativas e sua rápida difusão no ciberespaço. Para tanto, foi utilizada uma metodologia de caráter descritivo, fundamentada em legislações, doutrinas, literatura, artigos e casos concretos. A pesquisa conclui que a atual legislação brasileira é ineficaz ao tratar sobre o assunto em comparação a outros ordenamentos jurídicos, carecendo de revisão e complementação para mitigar a problemática que vitima tantos indivíduos, sobretudo mulheres e crianças, violando o princípio fundamental da dignidade da pessoa humana e o direito a honra e imagem.

Palavras-chave: Deepfakes; Pornografia Artificial Não Consensual; Redes Sociais; Responsabilidade; Legislação.

INTRODUÇÃO

A presente pesquisa discorre sobre as deepfakes feitas com conteúdo pornográfico não consensual, tema relevante no atual contexto social em razão do avanço cada vez maior de tecnologias de inteligência generativa e sua difusão no ciberespaço.

A deepfake é uma espécie de inteligência artificial generativa que funciona para criar imagens, vídeos e áudios manipulados, através de um banco de dados, sendo extremamente similares com imagens reais, com grande capacidade de serem interpretadas como verdadeiras.

Tal tecnologia generativa está cada vez mais avançada e difundida no mundo digital, com fácil acesso e uso para os usuários. Apesar de também serem usadas para criações artísticas úteis

¹ Discente do 5º Termo do curso de Direito do Centro Universitário Antônio Eufrásio de Toledo de Presidente Prudente. E-mail: maria.luiza@toledoprudente.edu.br.

ou interessantes, frequentemente são usadas para promover violência, caso da pornografia artificial não consensual.

Assim, diante dessa problemática, o artigo tem como objetivo questionar quais são os limites do uso dessa tecnologia, quais são os impactos e danos causados por ela e como a atual legislação trata a questão, nacional e globalmente. Com uma metodologia de caráter descritivo, pautada na análise bibliográfica, legislativa e de casos concretos, busca fomentar o debate sobre o tema com intuito de mitigar a criação e disseminação de deepfakes de conteúdo pornográfico não consensual.

1. PROBLEMÁTICA

As deepfakes surgiram em 2017, na rede social Reddit, quando os usuários passaram a criar montagens com teor sexual com sobreposições de imagens, vídeos e falas de outras pessoas.

Em janeiro de 2024 a cantora e compositora estadunidense Taylor Swift foi vítima do compartilhamento de deepfakes de teor pornográfico divulgadas em redes sociais, sobretudo no X, antigo Twitter, caso em que, em razão de sua influência, repercutiu massivamente, positiva e negativamente.

A título de ilustração, somente uma das imagens, dentre diversas postagens, obteve o número de 47 milhões de visualizações antes de ser removida, deixando claro o dano irreparável na imagem da cantora. Por outro lado, a viralização do caso gerou comoção social e diversos usuários se movimentaram em defesa da artista, criando a #ProtectTaylorSwift para encobrir as publicações, além da importante visibilidade dada ao fenômeno que já vitimava outras pessoas, celebridades ou não.

A influência da cantora e seus fãs serviu para pressionar fomentar o debate acerca das deepfakes de pornografia artificial não consensual para proteger os indivíduos contra violações de privacidade e imagem, e para que as fotos fossem removidas das redes sociais.

No entanto, com outras vítimas, anteriores e posteriores ao caso, celebridades menos influentes e pessoas desconhecidas, isso não ocorreu, permitindo a continuidade da exposição violenta e seus danos subsequentes, como ocorreu com Noelle Martin, ativista dos direitos das mulheres na era digital, que teve seu rosto utilizado sem seu consentimento em deepfakes pornográficas que ficaram em circulação por anos.

Uma pesquisa realizada em 2019 pela empresa holandesa de segurança cibernética, DeepTrace Labs, constatou que 96% dos vídeos online de deepfake eram de conteúdo pornográfico não consensual (Ajder et al, 2019). O estudo trouxe ainda outro dado alarmante: o número de visualizações nos vídeos dos quatro maiores sites dedicados a pornografia deepfake totalizou 134.364.438 (cento e trinta e quatro milhões, trezentos e sessenta e quatro mil e quatrocentas e trinta e oito), expondo o tamanho da comunidade público-alvo consumidora de tal conteúdo.

Diversas empresas, como Meta, Google e Microsoft desenvolveram para suas plataformas marcadores para conteúdos criados com inteligência artificial, no entanto, esses rótulos não são aplicados em todos os conteúdos artificiais, em razão de falhas de identificação das plataformas.

Não obstante, é necessário frisar que mesmo com rótulos de IA nas publicações de pornografia artificial, ou até quando sem, mas que os usuários conseguem facilmente identificar que aquele conteúdo não é real, tal fato não diminui as consequências para as vítimas. A título de exemplo, no caso da Taylor Swift, as imagens eram visivelmente falsas, e os usuários do X tinham conhecimento disso, dado que o termo “Taylor Swift AI” se tornou um *trending topics*² na plataforma, revelando que a rotulagem como conteúdo artificial seria insuficiente, pois não mudaria a intenção e os efeitos das postagens.

A autora inglesa Laura Bates, conhecida por suas obras na temática feminista, discorre sobre a misoginia online em seu livro “Homens que Odeiam Mulheres” (Bates, 2020, E-book, p. 172): “A ideia de que o abuso online é algo inofensivo que acontece em uma bolha separada, sem poder afetar sua vida real, é algo que você só pode realmente aceitar se nunca sofreu tal abuso.” (Em tradução livre). Como aponta a autora, mesmo que os ataques aconteçam no mundo online, as consequências são vivenciadas no mundo real.

As comunidades misóginas extremistas utilizam da manipulação digital para depreciar suas vítimas já sabendo que essas terão sequelas permanentes, e esse é o objetivo: atacá-las visceralmente com a ciência de que os danos são irreparáveis em suas vidas. Isso, pois uma vez que as imagens ou vídeos são publicados nas mídias sociais, mesmo que removidos posteriormente - o que frequentemente não acontece -, não há mais como controlar devidamente a propagação, nada impede que os usuários façam downloads dos arquivos, tirem prints ou outros atos similares.

² Termo utilizado para as temáticas em evidência, os assuntos virais em determinado momento na rede social X.

As vítimas de tal prática repugnante, reduzidas a um objeto sexual, são majoritariamente mulheres, mas também há materiais que retratam crianças e adolescentes sendo sexualmente exploradas (Saner, 2024). Essas, sofrem os impactos da fantasia alimentada pela cultura da misoginia e pedofilia, como a humilhação coletiva perante desconhecidos, amigos e familiares, o bullying na escola e no ambiente de trabalho, o assédio sexual e os traumas psicológicos.

2. RESPONSABILIDADE DAS PLATAFORMAS

Como mencionado anteriormente, determinadas plataformas possuem tecnologias para classificar publicações como conteúdos criados por inteligência artificial. No entanto, isso não basta para a pornografia artificial não consensual, já que o resultado é o mesmo, é necessário que tais conteúdos nem sequer tenham permissão para circulação online.

Os moderadores são os funcionários das plataformas responsáveis por analisar e controlar os conteúdos veiculados nas redes sociais, deletando se necessários aqueles que contenham violações das diretrizes de uso da rede.

A rede social do Elon Musk, X, sempre foi notavelmente conhecida por sua falha moderação. Com a repercussão do caso de Taylor, e a pressão dos usuários, a rede X afirmou que estaria "removendo ativamente todas as imagens identificadas e tomando as medidas apropriadas contra as contas responsáveis por publicá-las" (G1, 2024). Todavia, após adquirir o antigo Twitter em 2022, Elon decidiu pela demissão em massa de funcionários de diversas áreas (VEJA, 2022), inclusive responsáveis pela moderação de conteúdo, sob alegação de incentivar a liberdade de expressão, sendo claramente conivente com os casos de violência sexual e outras violações de direitos humanos na plataforma, que já aconteciam muito antes de vitimar a cantora.

Na medida em que a inteligência artificial conhecidamente não possui parâmetros éticos bem definidos, as redes sociais devem os estabelecer em suas diretrizes e funcionamento. As plataformas precisam garantir a segurança de seus usuários, não os expondo a conteúdos violentos que infrinjam direitos humanos fundamentais, adotando medidas preventivas, como a moderação de conteúdo eficaz e rápida, sob pena de serem responsabilizados solidariamente aos que publicam as postagens.

3. DAS LEGISLAÇÕES

Nos últimos anos, inúmeras legislações que discorrem sobre inteligência artificial, inclusive generativa, foram implementadas em todo globo.

A título de exemplo, a União Europeia criou em 2024 o AI Act (Lei de Inteligência Artificial), que aborda de forma abrangente o assunto, buscando regulamentar e proteger a população de violações de direitos fundamentais como a democracia e Estado de Direito realizadas por intermédio da inteligência artificial.

A lei europeia, em seu artigo 3º, § 60, conceitua a deepfake estruturando a tecnologia em quatro pontos: a) conteúdo de imagem, áudio ou vídeo; b) gerado ou manipulado por IA; c) que se assemelha a pessoas, objetos, lugares, entidades ou eventos existentes; d) que pareceria falsamente autêntico ou verdadeiro a uma pessoa. Dessa forma, é possível facilmente enquadrar publicações e conteúdos como deepfakes para alcançar, quando necessário, a responsabilização e sanções de determinadas condutas, caso das deepfakes de pornografia não consensual.

Além disso, a legislação em questão também classifica o uso de IA em graus de risco - como risco alto, por exemplo, no caso das deepfakes que podem causar violações de direitos de indivíduos ou de toda coletividade -, explicita quais são as práticas proibidas com o uso de IA, como as deepfakes de conteúdo malicioso, além de determinar obrigações de transparência como o dever de inserir marcações/rótulos em conteúdos gerados com IA, rastreabilidade e responsabilidade das deepfakes, bem como sanções para as condutas proibidas.

No Brasil, a Constituição Federal de 1988 pautada no princípio fundamental da dignidade da pessoa humana, preceitua o direito a intimidade, a vida privada, a honra e a imagem das pessoas.

No entanto, no atual ordenamento jurídico brasileiro, a legislação acerca da IA e, especificamente, das deepfakes, não é tão clara e concisa quanto à europeia. No Marco Civil da Internet (Lei 12.965/2014), principal legislação sobre o tema atualmente, é clara a preocupação do legislador com os direitos constitucionais mencionados, entretanto, não postula efetivamente sobre a responsabilização e sanções para mitigar tais condutas nefastas.

O artigo 2º da referida legislação estabelece a liberdade de expressão como fundamento. Contudo, tal fundamento conhecidamente possui restrições, os usuários devem ter o poder de se expressar livremente, salvo quando violam direitos de terceiros e/ou diretrizes de uso das redes sociais. Acerca da questão, o autor Vitor Hugo Pereira Gonçalves, presidente do Instituto SIGILO, discorre (Gonçalves, 2017, E-book, p. 6):

Para o Marco Civil, a internet é a nova *Ágora* grega ou *Fórum Romano*, uma praça virtual que reúne a todos que queiram se manifestar sobre a pólis ou o Estado. É o lugar da manifestação e da liberdade. A liberdade de expressão na internet, nesse sentido, é a dimensão extrínseca da democracia digital. É o princípio de afirmação de todas as cidadanias reunidas em participação direta, o cidadão “total”.

Para tanto, o Marco Civil deve garantir e disciplinar o uso dessa nova *ágora*. A liberdade de expressão é o fundamento, o princípio, mas os incisos deste art. 2º são os limites e problematizam o a situação do cidadão “total”. Bobbio já alertou que pior que não ter democracia é o excesso dela. Qualquer princípio por natureza não é absoluto e tem os seus limites. Contudo, a liberdade de expressão, como teoria e prática nas redes de informação e comunicação, possui infinitas limitações que não são enfrentadas pelo Marco Civil.

Como o jurista menciona, a liberdade de expressão não é absoluta e deve respeitar outros direitos e princípios constitucionais, como o princípio da dignidade da pessoa humana e o direito a imagem e a honra.

Nesse sentido, consta claro o exemplo do artigo 3º, que traz os princípios do uso da internet, como os incisos V³ e VI⁴, que mencionam o estímulo de boas práticas e a responsabilização dos agentes por suas condutas. Além do art. 7º, que dispõe sobre os direitos dos usuários na internet como o direito a inviolabilidade da intimidade e da vida privada e sua proteção, bem como o direito a indenização pelos danos causados em consequência da violação destes⁵.

Apesar disso, a mesma lei estatui em seu texto limitações para a responsabilidade de tais condutas danosas ao ciberespaço:

Art. 18. O provedor de conexão à internet não será responsabilizado civilmente por danos decorrentes de conteúdo gerado por terceiros.

Art. 21. O provedor de aplicações de internet que disponibilize conteúdo gerado por terceiros será responsabilizado subsidiariamente pela violação da intimidade decorrente da divulgação, sem autorização de seus participantes, de imagens, de vídeos ou de outros materiais contendo

³ Art. 3º, V - preservação da estabilidade, segurança e funcionalidade da rede, por meio de medidas técnicas compatíveis com os padrões internacionais e pelo estímulo ao uso de boas práticas; (Lei 12.965/2014).

⁴ Art. 3º, VI - responsabilização dos agentes de acordo com suas atividades, nos termos da lei; (Lei 12.965/2014).

⁵ Art. 7º, I - inviolabilidade da intimidade e da vida privada, sua proteção e indenização pelo dano material ou moral decorrente de sua violação; (Lei 12.965/2014).

cenas de nudez ou de atos sexuais de caráter privado quando, após o recebimento de notificação pelo participante ou seu representante legal, deixar de promover, de forma diligente, no âmbito e nos limites técnicos do seu serviço, a indisponibilização desse conteúdo.

Parágrafo único. A notificação prevista no caput deverá conter, sob pena de nulidade, elementos que permitam a identificação específica do material apontado como violador da intimidade do participante e a verificação da legitimidade para apresentação do pedido.

Como destacado, a norma restringe as empresas provedoras de internet da responsabilização pelos conteúdos de pornografia artificial divulgados por terceiros, salvo nos casos em que após o recebimento da notificação do detentor do direito violado, não promove a indisponibilização do conteúdo.

Para mais, o Código Civil (Lei 10.406/2002), dispõe no artigo 20 sobre o direito dos indivíduos à proibição da utilização de imagens suas que possam lhe ferir a honra ou a respeitabilidade ⁶, situações essas enquadradas na disseminação de deepfakes de caráter pornográfico não consensual, por ferirem a honra objetiva das vítimas através da deterioração e sabotagem de imagens pessoais, disseminadas no ciberespaço para prejudicar suas reputações.

Ainda, o Código Penal introduziu em 2018 um novo dispositivo para tratar sobre as deepfakes de caráter pornográfico não consensual:

Art. 216-B. Produzir, fotografar, filmar ou registrar, por qualquer meio, conteúdo com cena de nudez ou ato sexual ou libidinoso de caráter íntimo e privado sem autorização dos participantes: (Incluído pela Lei nº 13.772, de 2018)

Pena - detenção, de 6 (seis) meses a 1 (um) ano, e multa.

Parágrafo único. Na mesma pena incorre quem realiza montagem em fotografia, vídeo, áudio ou qualquer outro registro com o fim de incluir pessoa em cena de nudez ou ato sexual ou libidinoso de caráter íntimo. (Incluído pela Lei nº 13.772, de 2018)

⁶ Art. 20. Salvo se autorizadas, ou se necessárias à administração da justiça ou à manutenção da ordem pública, a divulgação de escritos, a transmissão da palavra, ou a publicação, a exposição ou a utilização da imagem de uma pessoa poderão ser proibidas, a seu requerimento e sem prejuízo da indenização que couber, se lhe atingirem a honra, a boa fama ou a respeitabilidade, ou se se destinarem a fins comerciais. (Lei 10.406/2002).

Essa tipificação representa um grande avanço para a responsabilização dos responsáveis e proteção das vítimas no que concerne a criação de conteúdos que distorcem a realidade de sua imagem, ocasionando danos inestimáveis as suas reputações perante a sociedade.

No entanto, apesar das normas mencionadas, consta claro que a atual legislação brasileira não lida efetivamente com a responsabilização solidária dos provedores de internet, nem mencionada o dever de uma moderação prévio dos conteúdos publicados nas redes sociais.

CONSIDERAÇÕES FINAIS

Em conclusão, com base no exposto, constatou-se que a proteção contra deepfakes de caráter pornográfico não consensual ainda carece de regulamentação no atual ordenamento jurídico brasileiro, sendo necessário a inclusão de normas que lidem efetivamente com a responsabilização não apenas dos usuários que criam e publicam as deepfakes de pornografia não consensual, mas também, solidariamente, das empresas provedoras de conexão e plataformas que permitem que tais conteúdos fraudulentos sejam veiculados nas redes sociais por meio de uma moderação conivente com a violação dos direitos da personalidade e do princípio fundamental da dignidade da pessoa humana.

Assim, é necessário fomentar o debate sobre a questão, para estimular a criação de novos projetos de lei que tratem sobre a problemática e, posteriormente, normas jurídicas que alcancem efetivamente a responsabilização de todos os agentes responsáveis pela perpetuação desta violência que só avança com o desenvolvimento das tecnologias de inteligência artificial generativa, com intuito de mitigar a ocorrência de novos casos e proteger as vítimas e usuários no ciberespaço.

REFERÊNCIAS

AJDER, Henry. PATRINI, Giorgio. CAVALLI, Francesco. CULLEN, Laurence. **The State of Deepfakes: Landscape, Threats, and Impact**. Amsterdã, Deeptrace, 2019. Disponível em: https://regmedia.co.uk/2019/10/08/deepfake_report.pdf. Acesso em: 09 abr. 2025.

BATES, Laura. **Men Who Hate Women: From incels to pickup artists, the truth about extreme misogyny and how it affects us all**. Simon & Schuster UK, 2020.

BRASIL. **Constituição da República Federativa do Brasil de 1988**. Brasília, DF: Presidência da República. Disponível em: https://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 19 ago. 2025.

BRASIL. **Decreto-Lei nº 2.848, de 7 de dezembro de 1940.** Código Penal. Brasília: Presidência da República, 1940. Disponível em: https://www.planalto.gov.br/ccivil_03/decreto-lei/de12848compilado.htm. Acesso em 22 ago. 2025.

BRASIL. **Lei nº 10.406, de 10 de janeiro de 2002.** Institui o Código Civil. Brasília, DF: Presidência da República. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/2002/110406compilada.htm. Acesso em: 22 ago. 2025.

BRASIL. **Lei nº 12.965, de 23 de abril de 2014.** Estabelece princípios, garantias, direitos e deveres para o uso da Internet no Brasil. Brasília: Presidência da República. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/112965.htm. Acesso em: 19 ago. 2025.

CAMARGO, João Pedro Felix Ortiz. **Deepfake no Caso Taylor Swift: a inteligência artificial generativa e o cerceamento da segurança das mulheres no espaço digital.** Barbacena: Intercom, 2024. Disponível em: <https://sistemas.intercom.org.br/pdf/submissao/regional/13/2047/041820242355526621dd38c4c4f.pdf>. Acesso em: 09 abr. 2025.

DECCACHE, Matheus. **Elon Musk demite equipe do Twitter responsável por combate à desinformação.** São Paulo: VEJA, 2022. Disponível em: <https://veja.abril.com.br/mundo/elon-musk-demite-equipe-do-twitter-responsavel-por-combate-a-desinformacao/>. Acesso em: 09 abr. 2025.

GONÇALVES, Victor Hugo Pereira. **Marco Civil da Internet Comentado.** 1ª Edição, 2017. Rio de Janeiro: Atlas, 2016. ISBN 9788597009514. Disponível em: <https://integrada.minhabiblioteca.com.br/reader/books/9788597009514/>. Acesso em: 22 ago. 2025.

MATOS, Brenda Caires; SOTERO, Ana Paula da Silva. **DEEPPAKES E O DIREITO À IMAGEM NO CIBERESPAÇO: NECESSIDADE DE UMA NORMATIZAÇÃO ESPECÍFICA.** Cenas Educacionais, [S. l.], v. 7, p. e20107, 2024. DOI: 10.5281/zenodo.13772486. Disponível em: <https://revistas.uneb.br/cenaseducacionais/article/view/20107>. Acesso em: 20 ago. 2025.

MARTIN, Noelle. **Online predators spread fake porn of me. Here's how I fought back.** TEDxPerth, 2017. Disponível em: https://www.ted.com/talks/noelle_martin_online_predators_spread_fake_porn_of_me_here_s_how_i_fought_back/transcript. Acesso em: 07 abr. 2025.

MAZZEI, Amanda. **Caso Taylor Swift: sem regulamentação, IA pode ser ferramenta para agredir mulheres e violar intimidade.** São Paulo: CBN, 2024. Disponível em: <https://cbn.globo.com/tecnologia/noticia/2024/01/31/caso-taylor-swift-sem-regulamentacao-ia-pode-ser-ferramenta-para-agredir-mulheres-e-violar-intimidade.ghtml>. Acesso em: 09 abr. 2025.

PRESSE, France. **Imagens pornô falsas de Taylor Swift, geradas por inteligência artificial, causam indignação nos EUA.** São Paulo: G1, 2024. Disponível em: <https://g1.globo.com/pop>

<arte/musica/noticia/2024/01/26/imagens-porno-falsas-de-taylor-swift-geradas-por-inteligencia-artificial-causam-indignacao-nos-eua.ghtml>. Acesso em: 09 abr. 2025.

RIVELLI, Fabio. **Combatendo deepfakes: Desafios de gênero na regulação contra a Violência Digital**. São Paulo: Migalhas, 2024. Disponível em: <https://www.migalhas.com.br/coluna/ia-em-movimento/405037/combateendo-deepfakes>. Acesso em: 09 abr. 2025.

SANER, Emine. **Inside the Taylor Swift deepfake scandal: ‘It’s men telling a powerful woman to get back in her box’**. The Guardian, 2024. Disponível em: https://www.theguardian.com/technology/2024/jan/31/inside-the-taylor-swift-deepfake-scandal-its-men-telling-a-powerful-woman-to-get-back-in-her-box?_x_tr_sl=en&_x_tr_tl=pt&_x_tr_hl=pt&_x_tr_pto=tc. Acesso em: 08 abr. 2025.

SCHMIDT, Sarah. **Deepfakes, o novo estágio tecnológico da desinformação**. São Paulo: Pesquisa FAPESP, 2022. Disponível em: <https://revistapesquisa.fapesp.br/deepfakes-o-novo-estagio-tecnologico-das-noticias-falsas/>. Acesso em: 09 abr. 2025.

UNIÃO EUROPEIA. **EU Artificial Intelligence Act**. 2024. Disponível em: <https://artificialintelligenceact.eu/>. Acesso em: 20 ago. 2025.

WU, Rita. **O que imagens falsas de Taylor Swift indicam para o debate sobre uso de IA**. São Paulo: CNN, 2024. Disponível em: <https://www.cnnbrasil.com.br/blogs/rita-wu/tecnologia/o-que-imagens-falsas-de-taylor-swift-indicam-para-o-debate-sobre-uso-de-ia/>. Acesso em: 09 abr. 2025.