IMPLICAÇÕES LEGAIS E RISCOS EMERGENTES AO USO DA INTELIGÊNCIA ARTIFICIAL E A MANIPULAÇÃO DE IMAGENS: UMA ANÁLISE A PARTIR DE OBRAS CINEMATOGRÁFICAS

Ana Carolina Assumpção Paniagua dos Santos (IC) e Bruna Azzari Puga (Orientadora)

Apoio:PIVIC Mackenzie

RESUMO

A crescente utilização de ferramentas de inteligência artificial para a criação e a manipulação de imagens, sobretudo por meio de tecnologias generativas capazes de produzir conteúdos falsos com aparência realista, tem desafiado a eficácia das normas de proteção aos direitos da personalidade. Diante desse cenário, o presente estudo investigou em que medida tais ferramentas podem comprometer direitos fundamentais da população brasileira, em especial a proteção da imagem e da privacidade. Para tanto, adotou-se abordagem qualitativa, de natureza teórico-interpretativa, articulando as técnicas de análise bibliográfica, documental, normativa e cultural, esta última desenvolvida a partir do episódio *A Joan É Péssima*, da série *Black Mirror*, e da obra *Admirável Mundo Novo*, correlacionadas a casos concretos nacionais e internacionais. A pesquisa revelou que o ordenamento jurídico brasileiro, embora disponha de instrumentos relevantes, mostra-se insuficiente diante dos riscos específicos dessas tecnologias, sobretudo pela dificuldade de identificar, em tempo hábil, os usos ilícitos e pela fragilidade do consentimento formal do titular. Constatou-se, ainda, que experiências regulatórias estrangeiras, assentadas em deveres preventivos e em fiscalização especializada, oferecem parâmetros úteis ao debate nacional, hoje predominantemente reparatório. Concluiu-se que a proteção efetiva depende de uma governança equilibrada, que combine mecanismos de detecção, responsabilização proporcional e políticas educativas, de modo a conciliar inovação tecnológica, proteção dos direitos da personalidade e liberdade individual, sem convertê-la em instrumento de censura.

Palavras-chave: Inteligência Artificial. *Deepfakes*. Manipulação de Imagens.

ABSTRACT

The growing use of artificial intelligence tools for the creation and manipulation of images, particularly through generative technologies capable of producing false yet realistic content, has challenged the effectiveness of the legal rules protecting personality rights. Against this backdrop, this study examined the extent to which such tools may compromise the fundamental rights of the Brazilian population, especially the protection of one's image and privacy. To this end, a qualitative, theoretical-interpretative approach was adopted, combining bibliographical, documentary, statutory, and cultural analysis, the latter developed from the episode A Joan É Péssima, of the series Black Mirror, and the novel Admirável Mundo Novo (Brave New World), examined alongside concrete domestic and international cases. The research found that the Brazilian legal order, although equipped with relevant instruments, proves insufficient to address the specific risks posed by these technologies, chiefly owing to the difficulty of detecting unlawful uses in a timely manner and to the fragility of the data subject's formal consent. It further found that foreign regulatory experiences, grounded in preventive duties and specialised oversight, provide useful parameters for the national debate, which remains predominantly reparatory. The study concluded that effective protection depends on balanced governance, combining detection mechanisms, proportionate accountability, and educational policies, so as to reconcile technological innovation, the safeguarding of personality rights, and individual freedom, without turning the latter into an instrument of censorship.

Keywords: *Artificial Intelligence. Deepfakes. Image Manipulation*.

1. INTRODUÇÃO

A inteligência artificial ("IA") está cada vez mais presente no cotidiano da sociedade contemporânea, trazendo avanços em diferentes setores. No entanto, seu uso também levanta questões jurídicas, principalmente no que tange ao uso do direito de imagem de forma indevida, sem consentimento ou até mesmo a utilização para práticas ilegais.

Com a crescente utilização de ferramentas baseadas em IA para criação, manipulação e disseminação de imagens, como *ChatGPT* e *Copilot*, surgem novos riscos e desafios quanto à proteção da identidade, da imagem e da privacidade. Ademais, outra questão a ser enfrentada diz respeito à veracidade dos atos praticados, uma vez que, com o uso das tecnologias, torna-se desafiador para o público distinguir a autenticidade de alegações.

O direito à imagem constitui um dos elementos essenciais da personalidade jurídica, sendo expressamente tutelado pela Constituição Federal (1988) e pelo Código Civil Brasileiro (2002). Entretanto, o progresso de tecnologias como *deepfake*, tecnologia que utiliza redes artificiais e algoritmos de aprendizado de máquina para gerar conteúdos sintéticos realistas, algoritmos de reconhecimento facial e a criação de imagens sintéticas tem desafiado a eficácia das normas jurídicas vigentes, comprometendo a efetiva fiscalização e a punição das práticas ilícitas. Conforme reiterado pelo Supremo Tribunal Federal:

> Essas tecnologias representam riscos à saúde do debate público, a direitos individuais, como a honra, a imagem e o acesso a informações adequadas, e à normalidade de processos sociais sensíveis, como as eleições. (Supremo Tribunal Federal, 2024, p. 5).

A compreensão desse fenômeno exige situá-lo na própria trajetória da inteligência artificial, que se estende desde as investigações pioneiras de Alan Turing até as atuais aplicações em *deep learning*, subconjunto do aprendizado de máquina que viabiliza sistemas capazes de executar operações complexas em setores como saúde, transporte e, consequentemente, no próprio Judiciário (Cóbe *et al.*, 2020). Essas técnicas potencializam a automação em escala sem precedentes e ampliam a competitividade econômica, demandando, simultaneamente, políticas públicas capazes de acompanhar tais avanços (Cóbe *et al.*, 2020).

No âmbito jurídico, cresce a preocupação de que os algoritmos reproduzam uma lógica estritamente positivista, limitando a efetivação da justiça em casos de maior complexidade, o que reforça a necessidade de uma interpretação das normas capaz de integrar valores e direitos fundamentais às decisões (Siqueira; Wolowski, 2022). Discute-se, ainda, a responsabilidade civil decorrente do uso de sistemas automatizados, uma vez que a IA,

embora inovadora, dificulta a identificação do nexo causal e sugere a conveniência de regulamentação específica (Souza Antunes, 2019).

Nesse cenário, ganha relevo o fenômeno das *deepfakes*, que é capaz de gerar conteúdos audiovisuais com realismo suficiente para enganar o espectador. A técnica combina o *deep learning* com a manipulação de imagens, superando a mera edição para promover verdadeira alteração da realidade (Souza; Santaella, 2021).

Quando empregadas de forma ilícita, as *deepfakes* têm servido à disseminação de desinformação e à produção de conteúdo pornográfico sem o consentimento dos titulares, configurando grave violação dos direitos da personalidade, sobretudo do direito à imagem e à honra (Pinto; Oliveira, 2023). Nesse ponto, o Guia Ilustrado Contra as *Deepfakes*, elaborado pelo Supremo Tribunal Federal, evidencia de forma didática métodos de identificação desses conteúdos, apontando indícios como inconsistências na sincronia labial, variações na tonalidade da pele e demais anomalias sugestivas de adulteração (Supremo Tribunal Federal, 2024).

A manipulação de imagens, que outrora dependia de recursos manuais ou de ferramentas básicas, passa a apoiar-se em algoritmos sofisticados que ampliam a circulação de conteúdos fraudulentos e dificultam a identificação e a responsabilização dos agentes. Ao projetar os desdobramentos dessa transformação, Lee (2022) sustenta que a inteligência artificial não apenas aprimora o que já existe, mas transforma profundamente modos de vida, modelos de negócio e relações sociais, o que reforça a necessidade de referenciais jurídicos aptos a proteger os direitos individuais em um ambiente cada vez mais digitalizado. Embora dispositivos como o Código Civil (Brasil, 2002) e o Marco Civil da Internet (Brasil, 2014) ofereçam mecanismos de reparação de danos, a velocidade de disseminação dos conteúdos e o anonimato digital dificultam sua aplicação (Pinto; Oliveira, 2023), o que faz da autorregulação das plataformas uma alternativa relevante para a rápida remoção de conteúdos ilícitos (Supremo Tribunal Federal, 2024).

Essa análise ganha densidade ao ser confrontada com as experiências normativas internacionais. O relatório do European Parliamentary Research Service (2021, p. 24) adverte que a crescente utilização de *deepfakes* representa desafios significativos para a privacidade, a segurança digital e a integridade das informações disseminadas no ambiente virtual, o que permite estabelecer um paralelo entre as diretrizes da União Europeia e a legislação brasileira, a fim de identificar lacunas normativas e possíveis adaptações voltadas à proteção da imagem e da privacidade no contexto nacional.

Diante de um tema ainda emergente, as obras cinematográficas tornam-se importantes canais de comunicação, tendo em vista que estas são de fácil acesso a um

público massificado e permitem explorar diferentes realidades, servindo como instrumentos para análise e reflexão sobre o assunto.

Nesse sentido, destaca-se o primeiro episódio da sexta temporada da série *Black Mirror* (2011), *A Joan É Péssima*, que retrata de forma crítica a utilização da imagem e da privacidade de uma mulher comum para fins de entretenimento, sem que ela tenha consciência das eventuais consequências para sua vida pessoal e profissional. A narrativa expõe os impactos causados, mostrando como as novas tecnologias, utilizadas de forma prejudicial, podem distorcer a realidade e comprometer os direitos da personalidade. A análise será complementada pela obra *Admirável Mundo Novo* (Huxley, 2000), que aborda o controle social, a manipulação da realidade e os limites éticos do progresso tecnológico e a proporcionalidade em seu controle.

Dessa forma, a presente pesquisa propõe-se a analisar os impactos jurídicos do uso da inteligência artificial na manipulação de imagens, buscando compreender de que modo essa tecnologia pode comprometer os direitos fundamentais da sociedade, especialmente sobre a proteção da imagem, da privacidade e da liberdade.

Ademais, a investigação utilizará da análise de casos concretos, nacionais e internacionais, com o objetivo de identificar insuficiências e antecipar os possíveis rumos das decisões acerca do tema. Nesse contexto, a pesquisa busca responder: as ferramentas de inteligência artificial generativas, podem comprometer os direitos fundamentais, em particular a proteção da imagem e da privacidade? Assim, aprofundando o debate sobre as implicações legais dessas tecnologias, à luz da atual legislação nacional e das experiências sociais ilustradas em obras cinematográficas.

Para responder a tal indagação, este estudo adota abordagem qualitativa, de natureza teórico-interpretativa, articulando as técnicas de análise bibliográfica, documental, normativa e cultural, com raciocínio dedutivo que parte de princípios jurídicos e filosóficos para examinar suas implicações diante da inteligência artificial generativa e da manipulação de imagens. A investigação organiza-se em três eixos complementares:

(i) o eixo normativo-jurídico, dedicado à análise das normas constitucionais e infraconstitucionais de proteção à imagem, à honra e à privacidade, com destaque para o Código Civil, a Constituição Federal, a LGPD e o Marco Civil da Internet, bem como de relatórios e diretrizes institucionais, como o Guia do Supremo Tribunal Federal sobre *deepfakes* e o relatório do Parlamento Europeu sobre *deepfakes*;

(ii) o eixo bibliográfico e teórico, voltado à leitura crítica de doutrina e de artigos jurídicos e interdisciplinares sobre os impactos da IA nos direitos da personalidade, na responsabilidade digital e na ética tecnológica; e

(iii) o eixo cultural-interpretativo, que examina o episódio *A Joan É Péssima*, da série *Black Mirror*, e a obra *Admirável Mundo Novo*, correlacionando suas representações ficcionais com riscos jurídicos concretos.

A interpretação estrutura-se a partir de categorias construídas na revisão teórica: consentimento e autodeterminação informacional; exposição pública e vulnerabilidade; e veracidade e integridade da imagem, sistematizadas por meio de fichamentos e categorização temática, sob perspectiva crítica e constitucional, com ênfase na proteção da dignidade humana, da privacidade e da liberdade no ambiente digital, em observância às normas da ABNT e às diretrizes do edital PIVIC.

2. DESENVOLVIMENTO DO ARGUMENTO

A fim de demonstrar como os riscos abstratos identificados pela doutrina se materializam em situações concretas, a análise a seguir articula duas obras ficcionais: o episódio *A Joan É Péssima* e o romance *Admirável Mundo Novo*, com o caso CONAR/Volkswagen e a experiência regulatória britânica, evidenciando, a partir dessa convergência entre ficção e realidade, as lacunas do ordenamento brasileiro na tutela da imagem frente a tecnologias generativas.

2.1 A MANIPULAÇÃO DA REALIDADE EM "*A JOAN É PÉSSIMA*" E "*ADMIRÁVEL MUNDO NOVO*": Uma análise dos desafios jurídicos e filosóficos

A série antológica *Black Mirror* (2011), da Netflix, é uma produção de ficção científica que aborda os impactos sociais, éticos e psicológicos do avanço tecnológico na sociedade contemporânea. Cada episódio apresenta uma narrativa independente, explorando cenários distópicos onde o uso excessivo e/ou desregulado da tecnologia gera consequências negativas para as pessoas e o coletivo, propondo reflexões sobre a relação entre humanidade, tecnologia, inovação e poder, evidenciando os riscos associados à manipulação de dados e à perda da autonomia (Black Mirror, 2023).

O primeiro episódio da sexta temporada, lançado em 2023, *A Joan É Péssima*, assim como os demais da série, evidencia as consequências do uso desenfreado das tecnologias. A narrativa se inicia ao apresentar a rotina de Joan, uma mulher comum, que mantinha um relacionamento estável, um bom cargo profissional e certo reconhecimento social. Ao longo do episódio, observa-se o momento em que ela e seu companheiro acessam uma plataforma de streaming denominada "*Streamberry*", uma alusão à Netflix, onde se deparam com o

lançamento de uma série, que recebe o mesmo nome do episódio, retratando, em tempo real, os acontecimentos de sua própria vida, “interpretados” pela atriz Salma Hayek.

Posteriormente, descobre-se que a produção da série ocorre por meio de um sistema avançado de *deepfake*, capaz de reproduzir fielmente a imagem e os comportamentos da personagem principal, porém com a imagem da atriz Salma Hayek. A produção passa a expor situações íntimas da vida de Joan, incluindo conflitos pessoais, traições e sessões de terapia, todas retiradas de contexto e exibidas publicamente, violando sua privacidade. A exposição indevida transforma sua rotina em um verdadeiro caos, fazendo com que ela sofra rejeição social, perca seu relacionamento e seu emprego, além de ser alvo de julgamentos constantes.

Joan busca auxílio jurídico na tentativa de conter a situação. No entanto, sua advogada constata que, ao aceitar os termos e condições da plataforma, os quais não foram lidos previamente, a personagem autorizou o uso irrestrito de sua imagem e de seus dados pessoais pela empresa produtora. Dessa forma, verifica-se que, juridicamente, não haveria medidas capazes de impedir a continuação da exibição do seriado.

Além disso, o episódio mostra a própria atriz, Salma, buscando assessoria jurídica após perceber que sua imagem estava sendo associada a comportamentos considerados moralmente reprováveis. Contudo, assim como ocorreu com Joan, a atriz é informada de que seus direitos de imagem haviam sido igualmente cedidos à produtora, que poderia utilizá-los livremente com os atos praticados pela personagem.

Assim, a obra, ainda que ficcional, funciona como alerta para os riscos emergentes do uso inadequado da inteligência artificial e para a relativa facilidade com que normas de proteção de dados e de direitos de imagem podem ser contornadas. Assim, mostra-se a importância de uma análise prévia e aprofundada dos riscos, inclusive antes da formulação de novas normas protetivas para o ordenamento, como destaca Cristina Godoy Bernardo, ao afirmar que:

> Consequentemente, deve-se identificar as lacunas existentes no nosso ordenamento jurídico e as fraquezas de nossa Estratégia Brasileira de Inteligência Artificial para ser possível adotar medidas efetivas que nos permita edificar um plano de ações que nos coloque no plano internacional em uma posição competitiva, mesmo que seja em apenas uma área eleita como estratégica para alavancar o nosso avanço na era digital. (Oliveira, 2022, p. 3).

Diante do aperfeiçoamento das tecnologias de IA, ainda existem lacunas significativas na tutela dos direitos da personalidade, o que pode ameaçar direitos assegurados na Constituição Federal de 1988, como a inviolabilidade da honra, da imagem, da vida privada e

da intimidade, garantias concebidas em contexto muito distinto ao atual, quando não havia massificação de tecnologias capazes de reproduzir, manipular e difundir imagens em larga escala (Pinto; Oliveira, 2023).

Dentre os avanços promovidos pela LGPD, destaca-se o estabelecimento de princípios como a finalidade, a necessidade e a segurança no tratamento de dados pessoais (Brasil, 2018), os quais representam importantes instrumentos de controle frente ao uso de sistemas de IA. Todavia, conforme aponta Cristina Godoy Bernardo (Oliveira, 2022), a legislação ainda é insuficiente para enfrentar os riscos específicos de tecnologias mais avançadas, como sistemas de *deepfake* voltados à exploração da imagem, podendo ser observado, por exemplo, no episódio da série em que, mesmo após a violação de seu direito à privacidade, Joan não encontra amparo jurídico eficaz, uma vez que a própria plataforma havia se valido de uma lacuna que legitimava a atitude e inviabilizava qualquer medida capaz de fazê-la cessar.

Nesse sentido, embora a Lei Geral de Proteção de Dados ("LGPD") exija que o consentimento do titular seja livre e informado, não há mecanismos eficazes que assegurem, na prática, que os consumidores compreendam plenamente o alcance das autorizações concedidas. No mesmo contexto, Siqueira e Andrade (2024) ressaltam que, na era da pós-verdade, a *deepfake* compromete a própria verificação da veracidade dos conteúdos, o que reforça a importância de uma gestão proativa dos dados pessoais pelos usuários e da regulação das plataformas digitais como medidas de proteção da imagem.

Em diálogo com o episódio analisado*,* a obra distópica *Admirável Mundo Novo* (título original *Brave New World*), de Aldous Huxley, publicado em 1932, oferece chave complementar para análise, mostrando uma sociedade futurista, em que a ordem social é mantida por reproduções industrializadas, indivíduos pré-classificados em castas e condicionamento psicológico desde a infância, assegurando a conformidade e a obediência.

A narrativa apresenta recursos como engenharia biológica, hipnopedia (transmissão de informações a uma pessoa enquanto ela dorme, com o objetivo de influenciar o comportamento) e o uso de uma droga de controle social (Soma) para suprimir dissenso e singularidade, dramatizando como a evolução tecnológica, sem limites éticos, pode suprimir a autonomia individual em nome de uma "estabilidade coletiva" (Huxley, 2000).

De modo distinto, ainda que em certa medida semelhante à série da Netflix, Huxley apresenta dois pontos centrais: primeiro, como o uso desenfreado de tecnologias avançadas, sem limites éticos e/ou mecanismos de controle, pode alterar profundamente a percepção coletiva do aceitável; e segundo, como advertem Mulholland e Oliveira (2021), uma população pode ser moldada a acreditar naquilo que lhe é imposto como verdade, por meio de

tecnologias generativas, levando indivíduos à obediência a novas normas. No mesmo contexto, as *deepfakes*, ao se inserirem no debate público, tendem a aprofundar problemas já existentes de desinformação e polarização, ampliando o risco de que narrativas fabricadas sejam socialmente aceitas como verídicas (Mulholland; Oliveira, 2021).

Estudos empíricos recentes corroboram essa preocupação, ao evidenciar que a exposição a imagens produzidas por *deepfake* eleva a carga cognitiva e o viés de confirmação dos indivíduos, moldando a percepção acerca dos conteúdos manipulados (Amin; Hong; Mazhar, 2025).

Partindo desse pressuposto de que as pessoas podem ser induzidas a acreditar em narrativas que lhes são apresentadas, observa-se, por exemplo, no episódio que retrata a vida de Joan, a aceitação integral da ideia pelo público, sem qualquer reflexão crítica sobre a possibilidade de que tal situação, ocorrendo supostamente em tempo real, possa ser uma alegação falsa. Essa dinâmica se relaciona diretamente à Alegoria da Caverna, formulada por Platão no século IV a.C., na qual os homens, aprisionados, acreditam fielmente nas sombras projetadas diante de si, tomando-as como a realidade absoluta, sem buscar conhecer a verdade que se encontra além da caverna (Platão, 2018).

Nesse sentido, ao relacionar o episódio da série *Black Mirror*, percebe-se que os indivíduos, assim como os prisioneiros de Platão confundem sombras com o que é real, receberam o conteúdo sem questionar sua veracidade ou seus limites éticos, aceitando-o como realidade e, com isso, gerando consequências concretas na vida de uma pessoa. Tal reflexão leva à indagação sobre como tecnologias semelhantes podem ser inseridas na sociedade atual sem usos ilícitos, considerando o risco de criação de provas falsas ou acusações infundadas, motivadas por sentimentos hostis pessoais.

Assim, a obra funciona como um exemplo conceitual de reflexão: imagens e representações digitais podem operar, na atualidade, como instrumentos de condicionamento e legitimação, moldando percepções, normalizando comportamentos e dificultando a aplicação de mecanismos de vigilância (Souza; Santaella, 2021). Sob a perspectiva jurídica e filosófica, Huxley oferece, portanto, uma análise crítica que permite problematizar em que medida a sociedade estaria preparada para a utilização ética das novas tecnologias generativas e de que forma o poder público poderia exercer o controle sobre tais recursos, evitando tanto o uso abusivo da IA quanto a imposição de restrições que configurem invasão significativa da liberdade individual dos cidadãos.

Diante desse dilema, documentos internacionais oferecem relevantes parâmetros de governança: a UNESCO (2021) consagra os princípios da proporcionalidade e da supervisão humana como salvaguardas éticas, enquanto a Organização para a Cooperação e

Desenvolvimento Econômico (OCDE, 2019) recomenda que os sistemas de IA respeitem o Estado de Direito, a liberdade e a autonomia individual, mediante mecanismos apropriados de supervisão.

2.2 ANÁLISE DO CASO CONAR, ELIS REGINA E VOLKSWAGEN: "A vida imita a arte"

Como afirmou Oscar Wilde, "a vida imita a arte" (Wilde, 1891), portanto, é necessário cautela ao analisar obras literárias e cinematográficas, evitando tratá-las como meras representações distópicas sem conexão com a realidade. Desta forma, passa-se agora, após a análise das obras que abordam a manipulação de imagens e da realidade, a examinar casos concretos que evidenciam os desafios éticos e legais relacionados à inteligência artificial e à manipulação de imagens.

Em julho de 2023, o Conselho Nacional de Autorregulamentação Publicitária ("Conar") instaurou processo ético contra a Volkswagen em razão da utilização da imagem da cantora Elis Regina, falecida em 1982, em um comercial comemorativo dos 70 anos da marca. A peça publicitária utilizou a tecnologia de *deepfake* para recriar a figura da artista, que aparece ao lado de sua filha, Maria Rita, interpretando a canção "Como Nossos Pais", de Belchior, um dos maiores sucessos de sua carreira (CONAR, 2023a).

A denúncia apresentada após questionamentos de consumidores levantou duas questões centrais: (i) se o uso da imagem de Elis Regina foi ético e respeitoso; e (ii) se seria necessária a informação explícita de que o conteúdo havia sido gerado por IA. Em agosto, o colegiado decidiu arquivar o processo, entendendo que não houve desrespeito à figura da artista, uma vez que o uso de sua imagem ocorreu mediante consentimento dos herdeiros e retratando-a em atividade que realizava em vida.

Quanto à transparência sobre o uso da ferramenta, o Conar considerou que o princípio da transparência havia sido respeitado, já que a recriação era evidentemente de IA, sobretudo diante do conhecimento público sobre o falecimento da cantora (CONAR, 2023b). A doutrina especializada, ao analisar esse mesmo caso, reconhece que a reprodução digital da imagem, da voz e da performance de pessoa falecida constitui bem digital dotado de valor econômico, cuja utilização é admitida quando há consentimento dos titulares dos direitos, ressalvadas as hipóteses de disposição em sentido contrário (Chiarello; Herrera; Barbosa Filho, 2025).

Portanto, a decisão sugere que, enquanto o conteúdo não for prejudicial ou enganoso, e desde que haja consentimento dos detentores de direitos, o uso da tecnologia não é considerado problemático. Contudo, permanece a questão: como proceder em casos em que a manipulação da imagem é tamanha que não é possível sequer identificar o uso da

tecnologia? É justamente diante dessa dificuldade que experiências estrangeiras oferecem parâmetros

2.3 A EXPERIÊNCIA BRITÂNICA: Deveres preventivos e a regulação das *deepfakes* fora do Brasil

No Reino Unido, um avanço significativo ocorreu com a promulgação da Lei de Segurança Online (Online Safety Act – "OSA") em 2023. A legislação busca prevenir crimes cometidos online, incluindo a criminalização do compartilhamento de pornografia *deepfake* (United Kingdom, 2023). Entretanto, existem críticas de que apenas o criador da *deepfake* responderia pelo crime, enquanto outros envolvidos na disseminação do conteúdo permaneceriam impunes.

O fundamento da abordagem britânica repousa no chamado dever de cuidado (*duty of care*), princípio segundo o qual os provedores de serviços digitais, sobretudo as plataformas de compartilhamento de conteúdo entre usuários e os serviços de busca, assumem a obrigação positiva de identificar, avaliar e mitigar riscos antes que o dano se concretize (United Kingdom, 2023). Diferentemente de uma tutela meramente repressiva e centrada na reparação individual, a *Online Safety Act* adota uma lógica sistêmica e preventiva, exigindo das empresas avaliações de risco periódicas e a adoção de medidas proporcionais de moderação, sob fiscalização de autoridade reguladora independente com poder sancionatório.

Essa opção normativa aproxima o modelo britânico de uma responsabilização por deveres de conduta e o distingue da abordagem brasileira, ainda predominantemente voltada à reparação *ex post* do dano, contraste já ressaltado pela doutrina que compara o dever de cuidado britânico às propostas regulatórias nacionais (Conjur, 2025).

No campo específico de manipulação de imagens, a *Online Safety Act* qualificou o compartilhamento de pornografia *deepfake* não consentida como conduta criminalmente relevante, impondo às plataformas deveres proativos de prevenção e remoção, e não apenas a retirada do conteúdo após denúncia (United Kingdom, 2023).

Trata-se de resposta normativa direta ao mesmo risco ilustrado pelas obras analisadas neste estudo, em que a imagem alheia é explorada por meio de tecnologias generativas de fácil acesso. Cumpre observar, todavia, que o regime britânico foi construído em etapas, onde a *Online Safety Act* de 2023 tipificou, sobretudo, a conduta de compartilhar imagens íntimas não consentidas, incluindo *deepfakes*, deixando inicialmente em zona cinzenta a própria criação do material (United Kingdom, 2023).

Essa lacuna veio a ser progressivamente suprida: em janeiro de 2025, o Ministério da Justiça britânico anunciou a tipificação de um novo crime para a produção de *deepfakes*

sexualmente explícitos, de modo a abranger tanto quem cria quanto quem compartilha o conteúdo (United Kingdom, 2025b). Tal orientação consolidou-se com o *Data (Use and Access) Act* de 2025, que passou a criminalizar igualmente a criação de imagem íntima não consentida (United Kingdom, 2025a). Com isso, o ordenamento britânico ampliou a eficácia preventiva do regime, alcançando as duas pontas da conduta ilícita: a criação e a disseminação.

Persiste, ainda, a tensão entre proteção e liberdade que caracteriza toda regulação dessa natureza: as exigências de verificação de idade e de identidade, embora voltadas à tutela de grupos vulneráveis, transferem a entes privados o poder de coletar e gerir dados sensíveis dos usuários, suscitando questionamentos quanto à privacidade e à autodeterminação informacional. [1]

Em perspectiva comparada, a experiência do Reino Unido dialoga com o alerta do relatório europeu, para o qual a crescente utilização de *deepfakes* representa desafio significativo à privacidade, à segurança digital e à integridade informacional (European Parliamentary Research Service, 2021), e oferece parâmetro útil ao debate brasileiro.

Enquanto o Brasil ainda se ampara em instrumentos de reparação e em orientações institucionais de caráter didático, como o guia elaborado pelo Supremo Tribunal Federal (Supremo Tribunal Federal, 2024), o modelo britânico evidencia a viabilidade de uma regulação assentada em deveres preventivos e em fiscalização especializada, desde que acompanhada de salvaguardas contra o uso desproporcional do poder regulatório.

Esse exemplo demonstra que, embora haja avanços nas regulações ao redor do mundo, os governos ainda enfrentam desafios para equilibrar inovação tecnológica, proteção de direitos fundamentais e segurança digital. O debate internacional evidencia que o Brasil, ao construir sua própria regulamentação, deverá considerar não apenas os aspectos reparatórios, mas também os deveres preventivos e os impactos éticos da manipulação de imagens por inteligência artificial.

3. CONSIDERAÇÕES FINAIS

Este estudo teve por objetivo investigar em que medida as ferramentas generativas de inteligência artificial, sobretudo as voltadas à manipulação de imagens, podem comprometer direitos fundamentais da população brasileira, em especial a proteção da imagem e da

[1] No Brasil, o Estatuto Digital da Criança e do Adolescente, "ECA Digital" (Lei nº 15.211/2025) também veda a autodeclaração como mecanismo único de verificação de idade para produtos ou serviços impróprios a menores de 18 anos, exigindo mecanismos "confiáveis" de aferição a cada acesso. Contudo, em vez de deixar em aberto o destino dos dados coletados, a lei estabelece expressamente uma limitação de finalidade: os dados coletados especificamente para a verificação etária somente podem ser utilizados para esse fim, sendo vedada a criação de perfis comportamentais de crianças e adolescentes a partir dessas informações, inclusive para direcionamento de publicidade.

privacidade. Para tanto, adotou-se abordagem qualitativa e teórico-interpretativa, articulando a análise normativa, doutrinária e cultural, esta última desenvolvida a partir do episódio *A Joan É Péssima*, da série *Black Mirror*, e da obra *Admirável Mundo Novo*.

No eixo cultural-interpretativo, a análise conjunta das duas obras evidenciou como as tecnologias generativas podem operar como instrumentos de condicionamento e legitimação, moldando percepções e dificultando o controle sobre o uso da imagem alheia. O episódio protagonizado por Joan ilustrou a fragilidade do consentimento formal diante de termos de uso não compreendidos, ao passo que *Admirável Mundo Novo*, em diálogo com a Alegoria da Caverna, reforçou o risco de uma aceitação acrítica de narrativas fabricadas.

No eixo normativo-jurídico interno, o caso CONAR/Elis Regina demonstrou que, no Brasil, o emprego de *deepfakes* tende a ser admitido quando há consentimento dos titulares e ausência de conteúdo enganoso ou prejudicial, subsistindo, contudo, a dificuldade de tutela quando a manipulação é de tal ordem que impede a própria identificação da tecnologia empregada. Constatou-se, ainda, que instrumentos como o Código Civil, o Marco Civil da Internet e a Lei Geral de Proteção de Dados, embora relevantes, revelam-se insuficientes diante dos riscos específicos das tecnologias generativas.

Propõe-se, nesse contexto, a releitura do consentimento previsto no art. 7º, I, da LGPD à luz do princípio da boa-fé objetiva e da função social do contrato (arts. 113 e 421 do CC/2002): cláusulas genéricas de cessão irrestrita de imagem em termos de uso, quando não precedidas de informação específica sobre o emprego de tecnologias generativas, devem ser reputadas abusivas, por configurarem renúncia antecipada a direito da personalidade sem delimitação de finalidade.

Complementarmente, sustenta-se a viabilidade de se atribuir às plataformas digitais um dever de cuidado (*duty of care*) análogo ao modelo britânico, fundado na cláusula geral de responsabilidade civil do art. 927, parágrafo único, do CC/2002, que já contempla a responsabilização independente de culpa para atividades de risco, categoria na qual a operação de sistemas generativos de larga escala pode ser enquadrada.

No eixo comparado, a experiência do Reino Unido ofereceu parâmetro relevante: assentada no dever de cuidado (*duty of care*), a *Online Safety Act* de 2023 adotou lógica preventiva e sistêmica, posteriormente aperfeiçoada, quanto à criação de *deepfakes*, pelo *Data (Use and Access) Act* de 2025, de modo a alcançar tanto a produção quanto a disseminação do conteúdo ilícito. O contraste com o modelo brasileiro, ainda predominantemente reparatório, sinaliza a conveniência de deveres preventivos e de fiscalização especializada, acompanhados de salvaguardas contra o uso desproporcional do poder regulatório.

Desse percurso, sobressaem duas fragilidades estruturais. A primeira, técnica e operacional, reside na dificuldade de identificar, em tempo hábil, os usos ilícitos da IA sendo a detecção condição necessária para qualquer medida preventiva ou reparatória eficaz (Sensity, 2019; European Parliamentary Research Service, 2021). A segunda, normativa e ética, consiste em proteger a população sem comprometer princípios fundamentais, o que exige limites proporcionais, sob pena de converter mecanismos de proteção em instrumentos de censura.

Não se trata de rejeitar a tecnologia. A IA oferece ganhos reais à atividade profissional, à pesquisa e ao acesso à informação. O ponto crítico é a sua governança: é preciso combinar sua regulação (mecanismos de detecção), responsabilização jurídica (sanções proporcionais a autores e disseminadores de conteúdo ilícito) e políticas públicas educativas que fortaleçam uma sociedade informada, evitando que episódios como o apresentado na série Black Mirror ocorram num futuro possível.

Diante disso, cabe aos legisladores que não apenas elaborem normas que assegurem proteção legal, mas também que garantam sua aplicação prática, respeitando os limites éticos e constitucionais. O desafio global é evoluir não apenas na criação de tecnologias que facilitem a vida cotidiana, mas também na construção de mecanismos que protejam os cidadãos contra seus usos de má-fé. Somente assim será possível existir um futuro tecnológico seguro, em que inovação e direitos fundamentais coexistam de forma equilibrada.

4. REFERÊNCIAS

AMIN, Arslan; HONG, Yu; MAZHAR, Bilal. ***The influence of social media deepfake images on political ideology and polarization: the mediating roles of cognitive load and confirmation bias***. *Journal of Visual Literacy*, [s. l.], 2025. DOI: 10.1080/1051144X.2025.2543689. Disponível em: https://doi.org/10.1080/1051144X.2025.2543689. Acesso em: 14 ago. 2026.

BLACK MIRROR. **A Joan É Péssima**. Direção: Ally Pankiw. Produção: Charlie Brooker; Annabel Jones. Intérpretes: Annie Murphy; Salma Hayek Pinault; Michael Cera et al. Roteiro: Charlie Brooker [S. l.]: Netflix, 2023. Temporada 6, episódio 1 (61 min).

BRASIL. [Constituição (1988)]. **Constituição da República Federativa do Brasil de 1988**. Brasília, DF: Senado Federal, 1988. Disponível em: https://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 6 abr. 2025.

BRASIL. **Lei nº 10.406, de 10 de janeiro de 2002**. Institui o Código Civil. Brasília, DF: Presidência da República, 2002. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/2002/l10406compilada.htm. Acesso em: 6 abr. 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, DF: Presidência da República, 2014. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l12965.htm. Acesso em: 21 abr. 2025.

BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília, DF: Presidência da República, 2018. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm. Acesso em: 6 abr. 2025.

BRASIL. **Lei nº 15.211, de 17 de setembro de 2025**. Dispõe sobre a proteção de crianças e adolescentes em ambientes digitais (Estatuto Digital da Criança e do Adolescente). Diário Oficial da União: seção 1, edição extra-A, Brasília, DF, p. 1, 17 set. 2025. Disponível em: https://www2.camara.leg.br/legin/fed/lei/2025/lei-15211-17-setembro-2025-797997-norma-pl.html. Acesso em: 22 ago. 2026.

CHIARELLO, Felipe; HERRERA, Elon Caropreso; BARBOSA FILHO, Fernando Franco. **A tutela póstuma dos bens digitais e a herança digital: o debate sobre a (in)transmissibilidade de bens existenciais/personalíssimos**. Revista do Instituto Histórico e Geográfico Brasileiro, Rio de Janeiro, v. 186, n. 499, p. 190-221, 2025.

CÓBE, R. M. O.; NONATO, L. G.; NOVAES, S. F.; ZIEBARTH, J. A. **Rumo a uma política de Estado para inteligência artificial**. Revista USP, São Paulo, n. 124, p. 37-48, 2020. DOI: 10.11606/issn.2316-9036.v0i124p37-48.

CONAR – Conselho Nacional de Autorregulamentação Publicitária. **Processo ético: caso Volkswagen/Elis Regina**. São Paulo: Conar, 2023a. Disponível em: https://www.conar.org.br/. Acesso em: 4 fev. 2026.

CONAR – Conselho Nacional de Autorregulamentação Publicitária. **Processo ético nº 134/23: representação contra a Volkswagen do Brasil e a agência AlmapBBDO**. São Paulo: Conar, 2023b. Disponível em: http://www.conar.org.br/processos/detcaso.php?id=6354. Acesso em: 26 abr. 2025.

CONJUR. **O dever de cuidado como pilar da confiança digital.** [S. l.], 2025. Disponível em: https://www.conjur.com.br/2025-set-02/o-dever-de-cuidado-como-pilar-da-confianca-digital-comparando-a-legislacao-do-reino-unido-e-a-proposta-brasileira/. Acesso em: 21 abr. 2025.

EUROPEAN PARLIAMENTARY RESEARCH SERVICE. ***Tackling deepfakes in European policy***. *Estrasburgo: European Parliament*, 2021. Disponível em: https://www.europarl.europa.eu/stoa/en/document/EPRS_STU(2021)690039. Acesso em: 21 abr. 2025.

HUXLEY, Aldous. **Admirável Mundo Novo**. [S. l.]: Editora Globo, 2000.

LEE, Kai-Fu. **2041: como a inteligência artificial vai mudar sua vida nas próximas décadas**. [S. l.]: FisicalBook, 2022.

MULHOLLAND, Caitlin; OLIVEIRA, Samuel Rodrigues de. **Uma nova cara para a política? Considerações sobre *deepfakes* e democracia**. Direito Público, Brasília, v. 18, n. 99, p. 368-396, jul./set. 2021. DOI: 10.11117/rdp.v18i99.5773.

OCDE – Organização para a Cooperação e Desenvolvimento Econômico. ***Recommendation of the Council on Artificial Intelligence:* OECD/LEGAL/0449**. Paris: OCDE, 2019 (alterada em 2024). Disponível em: https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449. Acesso em: 14 ago. 2026.

OLIVEIRA, Cristina Godoy Bernardo de. **Desafios da regulação do digital e da inteligência artificial no Brasil**. Revista USP, São Paulo, n. 135, p. 137-162, 2022. DOI: 10.11606/issn.2316-9036.i135p137-162.

PINTO, Felipe Chiarello de Souza; OLIVEIRA, Gabriela Franklin de. **Não acredite em tudo que vê: *deepfake* pornography e responsabilidade civil no ordenamento jurídico brasileiro**. Revista Direito e Política, [s. l.], v. 18, n. 2, p. 427-451, 2023. DOI: 10.14210/rdp.v18n2.p427-451.

PLATÃO. **A República**. Tradução e organização de J. Guinsburg. 2. ed. São Paulo: Perspectiva, 2018.

SENSITY. ***The state of deepfakes: landscape, threats, and impact***. Amsterdã: Sensity, 2019. Disponível em: https://sensity.ai/reports/. Acesso em: 21 abr. 2025.

SIQUEIRA, D. P.; WOLOWSKI, M. R. de O. **Inteligência artificial e o positivismo jurídico: benefícios e obstáculos para efetivação da justiça**. Revista Brasileira de Direito, [s. l.], v. 18, n. 1, 2022.

SIQUEIRA, Mariana de; ANDRADE, Ester Jerônimo de. ***Deepfake* e privacidade: uma análise jurídica acerca da manipulação da imagem dos usuários**. Revista Foco, [s. l.], v. 17, n. 8, 2024. DOI: 10.54751/revistafoco.v17n8-014.

SOUZA, Carlos Eduardo de; SANTAELLA, Lucia. ***Deepfakes* na perspectiva da semiótica**. TECCOGS: Revista Digital de Tecnologias Cognitivas, São Paulo, n. 23, p. 26-44, 2021. DOI: 10.23925/1984-3585.2021i23p26-44.

SOUZA ANTUNES, H. **Inteligência artificial e a responsabilidade civil: enquadramento. Revista de Direito da Responsabilidade**, [s. l.], 2019.

SUPREMO TRIBUNAL FEDERAL (STF). **Guia ilustrado contra *deepfakes* 2024**. Brasília, DF: STF, 2024. Disponível em: https://portal.stf.jus.br/desinformacao/doc/Guia%20ilustrado%20Contra%20Deep%20Fakes_ebook%20(1).pdf. Acesso em: 2 abr. 2025.

UNESCO. ***Recommendation on the ethics of artificial intelligence***. Paris: Unesco, 2021. (SHS/BIO/REC-AIETHICS/2021). Disponível em: https://unesdoc.unesco.org/ark:/48223/pf0000380455. Acesso em: 14 ago. 2026.

UNITED KINGDOM. ***Data (Use and Access) Act* 2025**, Section 138. [S. l.]: Legislation.gov.uk, 2025a. Disponível em: https://www.legislation.gov.uk/ukpga/2025/18/section/138. Acesso em: 14 ago. 2026.

UNITED KINGDOM. *Ministry of Justice*. ***Government crackdown on explicit deepfakes***. GOV.UK, 7 jan. 2025b. Disponível em: https://www.gov.uk/government/news/government-crackdown-on-explicit-deepfakes. Acesso em: 14 ago. 2026.

UNITED KINGDOM. ***Online Safety Act* 2023**: Legislation.gov.uk, 2023. Disponível em: https://www.legislation.gov.uk/ukpga/2023/50. Acesso em: 21 abr. 2025.

WILDE, Oscar. ***Intentions***. Londres: James R. Osgood, McIlvaine & Co., 1891.

Contatos: anacapaniagua@gmail.com e bruna.azzari@mackenzie.br.