Cognitio Juris
ISSN 2236-3009

RESPONSABILIDADE CIVIL PELA VIOLAÇÃO DO DIREITO À IMAGEM POR DEEPFAKES GERADOS POR INTELIGÊNCIA ARTIFICIAL: CRITÉRIOS PARA INDIVIDUALIZAR A IMPUTAÇÃO

CIVIL LIABILITY FOR VIOLATIONS OF IMAGE RIGHTS THROUGH AI-GENERATED DEEPFAKES: CRITERIA FOR INDIVIDUALIZING ATTRIBUTION

Artigo submetido em 17 de novembro de 2026
Artigo aprovado em 18 de novembro de 2026
Artigo publicado em 18 de novembro de 2026

Cognitio Juris
Volume 16 – Número 59 – 2026
ISSN 2236-3009

.

Autor:
Victor Alexandre Barbosa Marques[1]
Renato Luiz Gomez da Silva[2]

.

RESUMO: O avanço da inteligência artificial generativa ampliou a produção de imagens, vídeos e áudios sintéticos capazes de atribuir a pessoas reais falas, condutas ou situações inexistentes, intensificando riscos ao direito à imagem. Este artigo analisa os critérios

jurídicos aplicáveis à responsabilidade civil dos diferentes agentes envolvidos na criação e circulação de deepfakes no ordenamento brasileiro. Adota-se método dedutivo, abordagem qualitativa e pesquisa jurídico-dogmática, bibliográfica, documental e jurisprudencial, com exame da Constituição Federal, do Código Civil, do Código de Defesa do Consumidor, do Marco Civil da Internet, da Lei Geral de Proteção de Dados, dos Decretos nº 12.975/2026 e nº 12.976/2026, da jurisprudência do Superior Tribunal de Justiça e da tese fixada pelo Supremo Tribunal Federal nos Temas 987 e 533, aperfeiçoada em 2026. Os resultados indicam que a imputação não pode ser uniforme: deve considerar participação causal, conhecimento da ilicitude, dever jurídico de agir, capacidade de prevenção ou mitigação, proveito econômico, natureza do dano, consentimento e possibilidade de identificação e prova. O regime das plataformas varia conforme o tipo de conteúdo, distinguindo ofensas à honra, conteúdo íntimo, crimes graves e anúncios ou impulsionamentos pagos. Conclui-se que não há solidariedade automática entre todos os participantes da cadeia, embora ela possa decorrer de coautoria do ilícito ou de regimes específicos aplicáveis às plataformas.

Palavras-chave: Deepfake; Direito à imagem; Inteligência artificial; Responsabilidade civil; Direitos da personalidade.

ABSTRACT: The advancement of generative artificial intelligence has expanded the production of synthetic images, videos and audio capable of attributing to real people statements, conduct or situations that never occurred, increasing risks to image rights. This article analyzes the legal criteria applicable to the civil liability of the different actors involved in the creation and circulation of deepfakes under Brazilian law. It adopts a deductive method, a qualitative approach, and legal-dogmatic, bibliographic, documentary and case-law research, examining the Federal Constitution, the Civil Code, the Consumer Protection Code, the Brazilian Internet Civil Rights Framework, the General Data Protection Law, Decrees No. 12,975/2026 and No. 12,976/2026, Superior Court of Justice case law, and the Federal Supreme Court ruling in Themes 987 and 533, refined in 2026. The findings indicate that attribution cannot be uniform and must consider causal participation, knowledge of

unlawfulness, legal duty to act, capacity for prevention or mitigation, economic benefit, nature of harm, consent, and the possibility of identification and proof. Platform liability varies according to the type of content, distinguishing offenses against honor, intimate content, serious crimes, and paid advertising or boosting. The article concludes that there is no automatic joint liability among all participants in the technological chain, although joint liability may result from co-authorship of the unlawful act or from specific regimes applicable to platforms.

Keywords: Deepfake; Image rights; Artificial intelligence; Civil liability; Personality rights.

1 INTRODUÇÃO

Em 21 de maio de 2026, o Diário Oficial da União publicou o Decreto nº 12.976, que estabelece diretrizes para a proteção de mulheres no ambiente digital. Seu art. 3º, I, define conteúdo íntimo como imagem, vídeo, áudio ou texto que exponha nudez, seminudez, ato sexual ou contexto sexualizante, “ainda que produzido e manipulado, no todo ou em parte, por sistema de inteligência artificial ou recurso tecnológico equivalente” (Brasil, 2026b, art. 3º, I). A inclusão expressa do conteúdo sintético na definição normativa revela que o deepfake deixou de ser curiosidade tecnológica e passou a ser objeto de regulação específica. Resta saber quem responde quando ele causa dano.

A inteligência artificial generativa ampliou de forma expressiva a capacidade de criar e manipular imagens, vídeos e áudios. Entre suas aplicações estão os deepfakes, conteúdos sintéticos capazes de reproduzir rosto, voz, gestos e outros elementos identificadores de uma pessoa com elevado grau de verossimilhança. Quando utilizados sem consentimento, esses recursos podem atingir diretamente direitos da personalidade, especialmente a imagem, a honra, a privacidade e a identidade. A Constituição Federal protege a imagem no art. 5º, V e X, enquanto o Código Civil disciplina os direitos da personalidade e o dever de reparar decorrente de ato ilícito ou abuso de direito (Brasil, 1988, 2002).

A literatura recente já demonstra que o ordenamento brasileiro possui instrumentos para reagir a violações produzidas por inteligência artificial, mas ainda há dificuldade para distribuir a responsabilidade quando criação, publicação, impulsionamento, hospedagem e fornecimento da tecnologia envolvem agentes distintos (Alves; Sampaio Neto, 2025; Amaral et al., 2026). Não se pretende, portanto, apenas verificar se a vítima de deepfake possui tutela jurídica, mas identificar critérios capazes de diferenciar a posição do criador, do divulgador, da plataforma digital e do fornecedor do sistema de inteligência artificial.

Diante desse cenário, formula-se o seguinte problema de pesquisa: quais critérios jurídicos podem ser utilizados para definir a responsabilidade civil dos diferentes agentes envolvidos na criação e divulgação de deepfakes que violem o direito à imagem no ordenamento jurídico brasileiro? A questão exige considerar que a tecnologia pode ser utilizada em situações diversas, como conteúdo íntimo sintético, fraude, publicidade enganosa ou falsa atribuição de falas, e que o regime jurídico aplicável pode variar conforme a natureza do conteúdo e a participação de cada agente.

Parte-se da hipótese de que o sistema brasileiro contém fundamentos normativos capazes de assegurar tutela preventiva e reparatória, porém a atribuição de responsabilidade não pode ser uniforme. A análise deve considerar a contribuição causal de cada participante, o conhecimento da ilicitude, o dever jurídico de agir, a capacidade de prevenir ou mitigar o dano, o eventual proveito econômico, o contexto de consentimento e os elementos disponíveis para identificação e prova.

O objetivo geral é analisar os critérios jurídicos aplicáveis à responsabilidade civil dos diferentes agentes envolvidos na criação e divulgação de deepfakes que violem o direito à imagem no ordenamento jurídico brasileiro. Para alcançar esse objetivo, estabelecem-se os seguintes objetivos específicos:

a) examinar a proteção constitucional e civil conferida ao direito à imagem;

b) compreender o fenômeno dos deepfakes e suas principais formas de lesão;

c) analisar os pressupostos da responsabilidade civil aplicáveis aos danos decorrentes da utilização de deepfakes;

d) identificar as posições jurídicas do criador, do divulgador, da plataforma digital e do fornecedor de sistemas de inteligência artificial;

e) examinar a jurisprudência do Superior Tribunal de Justiça sobre o direito à imagem e dois julgados estaduais sobre deepfakes, estes conhecidos pela descrição feita na produção acadêmica, observando as dificuldades de imputação da responsabilidade e de reparação dos danos.

Metodologicamente, adota-se o método dedutivo, com abordagem qualitativa e pesquisa jurídico-dogmática, apoiada em revisão bibliográfica, análise documental e exame jurisprudencial. O detalhamento das fontes, dos critérios de seleção, dos eixos de leitura e do recorte temporal é apresentado na seção 3.

O artigo está organizado em três núcleos após esta introdução. O primeiro apresenta o referencial teórico sobre direitos da personalidade, direito à imagem, inteligência artificial generativa e deepfakes, e delimita a lacuna da pesquisa. O segundo explicita o método. O terceiro concentra os resultados e a discussão, examinando os pressupostos da responsabilidade civil, a posição dos diferentes agentes, os regimes diferenciados das plataformas e as atualizações normativas e jurisprudenciais de 2026, sistematiza os critérios de imputação e enfrenta as principais objeções. Ao final, as considerações finais respondem diretamente ao problema proposto.

2 REFERENCIAL TEÓRICO

2.1 Direitos da personalidade e proteção constitucional da imagem

Os direitos da personalidade protegem atributos essenciais da pessoa e se relacionam diretamente com a dignidade humana. Schreiber (2014) destaca que sua tutela deve acompanhar as transformações sociais e tecnológicas, porque novas formas de exposição podem atingir dimensões centrais da autonomia pessoal. No ordenamento brasileiro, essa proteção possui fundamento constitucional e civil: o art. 1º, III, consagra a dignidade da pessoa humana, e o art. 5º, V e X, assegura tutela da honra, da intimidade, da vida privada e da imagem, além da indenização correspondente quando houver violação (Brasil, 1988).

No Código Civil, os arts. 11 a 21 integram o regime dos direitos da personalidade. O art. 12 admite providências para fazer cessar ameaça ou lesão, sem prejuízo de perdas e danos; o art. 20 disciplina a utilização da imagem e da palavra em hipóteses capazes de atingir interesses personalíssimos ou explorar economicamente a pessoa; e o art. 21 protege a vida privada. Esses dispositivos demonstram que a tutela da imagem não se limita à indenização posterior, admitindo também medidas preventivas e inibitórias (Brasil, 2002).

A doutrina distingue imagem-retrato e imagem-atributo. A primeira se refere à representação identificável da pessoa, compreendendo aparência física e outros elementos exteriorizados; a segunda corresponde à projeção social da personalidade e à forma pela qual o indivíduo é reconhecido perante terceiros. Teffé (2017) observa que, na internet, a manipulação e a circulação da imagem tornam essa proteção particularmente complexa, pois um mesmo conteúdo pode atingir simultaneamente a representação física e a identidade social do titular.

A jurisprudência do Superior Tribunal de Justiça reforça a autonomia e a dimensão patrimonial do direito à imagem. No REsp 267.529/RJ, a Quarta Turma reconheceu que o direito à imagem possui conteúdo moral e patrimonial e que a utilização não autorizada com finalidade econômica gera dever de indenizar, entendimento posteriormente consolidado na Súmula 403, segundo a qual independe de prova do prejuízo a indenização pela publicação não autorizada de imagem com fins econômicos ou comerciais (Brasil, 2000, 2009). Essa

orientação é relevante para deepfakes empregados em publicidade, monetização ou falsa associação comercial.

2.2 Direito à imagem no ambiente digital e proteção de dados

A internet ampliou a capacidade de captar, copiar, armazenar, modificar e redistribuir imagens. Uma fotografia publicada em rede social pode ser reutilizada em contexto completamente diferente daquele inicialmente autorizado. Teffé (2017) sustenta que a exposição digital não elimina a autonomia do direito à imagem e que, em regra, o uso não autorizado deve ser analisado independentemente de eventual ofensa adicional à honra, ressalvadas as circunstâncias que legitimem a utilização.

Essa lógica se torna ainda mais importante diante da inteligência artificial. A disponibilização pública de uma fotografia não equivale a consentimento genérico para a criação de representações sintéticas. O consentimento deve ser interpretado conforme finalidade e contexto, pois a autorização para uma publicação específica não implica concordância com nudez artificial, declaração inexistente, publicidade não autorizada ou associação a fraude. A possibilidade de replicação também aumenta a extensão temporal do dano e dificulta sua contenção (Alves; Sampaio Neto, 2025).

O Marco Civil da Internet estabelece princípios, direitos e deveres aplicáveis ao ambiente digital, disciplina a guarda de registros e organiza a responsabilidade de provedores por conteúdo de terceiros. A Lei Geral de Proteção de Dados, por sua vez, protege direitos fundamentais de liberdade, privacidade e livre desenvolvimento da personalidade no tratamento de dados pessoais. Imagens, gravações de voz e outros elementos identificadores podem integrar operações de tratamento, conforme a finalidade e o contexto em que forem coletados, combinados ou utilizados (Brasil, 2014, 2018).

A LGPD não substitui a tutela civil da imagem nem resolve, isoladamente, toda controvérsia envolvendo deepfakes. Sua incidência deve ser articulada com os direitos da personalidade,

a responsabilidade civil e as regras de intermediação da internet. Essa leitura sistemática evita tanto a falsa ideia de ausência de proteção quanto a aplicação automática de um único diploma legal a situações tecnologicamente e juridicamente distintas.

2.3 Inteligência artificial generativa e conceito de deepfake

A inteligência artificial generativa compreende sistemas capazes de produzir novas saídas a partir de padrões aprendidos em grandes conjuntos de dados. No campo audiovisual, esses sistemas podem gerar imagens, sintetizar vozes, substituir rostos, alterar movimentos labiais e recombinar elementos de registros reais. O termo deepfake resulta da associação entre deep learning e fake e é empregado para designar conteúdos sintéticos ou manipulados com elevado grau de realismo (Amaral et al., 2026).

Do ponto de vista técnico, boa parte desses conteúdos é produzida por redes adversariais generativas (generative adversarial networks, GANs), arquitetura em que duas redes neurais são treinadas em oposição: uma geradora, que produz imagens ou sons sintéticos, e uma discriminadora, que tenta distinguir o material sintético do autêntico. A cada ciclo, a geradora aprende a enganar a discriminadora, e o resultado é um conteúdo progressivamente mais difícil de diferenciar do real (Westerlund, 2019). Esse desenho tem uma consequência jurídica relevante: as técnicas de detecção tendem a ser incorporadas ao próprio treinamento dos geradores, numa corrida em que cada avanço na identificação da fraude alimenta a sofisticação da fraude seguinte. Por isso, a literatura técnica não aposta em uma única solução e combina, como contramedidas, regulação, políticas das empresas, educação dos usuários e tecnologias de detecção e autenticação (Westerlund, 2019). Para o Direito, isso significa que a responsabilização não pode depender da premissa de que a falsificação será sempre tecnicamente demonstrável.

Os deepfakes não são ilícitos por natureza. Há usos consentidos em obras audiovisuais, educação, acessibilidade, publicidade e entretenimento. A qualificação jurídica depende da

finalidade, do consentimento, do contexto de circulação, do grau de autenticidade aparente e dos efeitos produzidos. Para o Direito, portanto, o ponto central não é a arquitetura computacional em si, mas a utilização da tecnologia para reconstruir ou atribuir artificialmente aspectos da identidade de uma pessoa real.

O aumento da verossimilhança reduz sinais perceptíveis de manipulação e pode dificultar a reação da vítima e de terceiros. Conteúdos sintéticos podem atribuir fala, comportamento, nudez, participação em fraude, apoio a produto ou posição pública inexistente. Nesses casos, a tecnologia potencializa formas tradicionais de uso indevido da imagem e cria dificuldades adicionais de prova, rastreabilidade e identificação do responsável (Ebling; Ribeiro, 2024; Amaral et al., 2026).

2.4 Deepfakes como forma de violação do direito à imagem

A principal particularidade do deepfake está na atribuição artificial de uma realidade ao titular da imagem. Em vez de apenas reproduzir um registro verdadeiro sem autorização, o conteúdo pode representar a pessoa em situação que nunca ocorreu. A manipulação de rosto e voz pode atingir a imagem-retrato e, simultaneamente, a imagem-atributo, especialmente quando atribui comportamento, declaração ou vínculo inexistente (Teffé, 2017; Amaral et al., 2026).

As formas de dano variam conforme a finalidade. Em fraudes, a imagem pode conferir credibilidade à operação e induzir terceiros a erro; em conteúdos íntimos sintéticos, corpo ou rosto podem ser artificialmente inseridos em nudez ou contexto sexual sem consentimento; em publicidade, a montagem pode criar falsa associação entre pessoa e produto; e, em ataques reputacionais, a tecnologia pode produzir declarações aparentemente autênticas. A diversidade de usos impede que todos os casos sejam enquadrados da mesma maneira.

O debate regulatório permanece em evolução. O Projeto de Lei nº 2.338/2023, aprovado no Senado e remetido à Câmara dos Deputados, continuava, em 12 de novembro de 2026,

aguardando parecer do relator na Comissão Especial sobre Inteligência Artificial. O projeto não constitui direito vigente, mas demonstra que o regime jurídico da inteligência artificial ainda está em construção legislativa e deve ser distinguido das normas já aplicáveis (Brasil, 2023). Em seu conteúdo, o projeto adota classificação dos sistemas por nível de risco e assegura aos afetados direitos de explicação, contestação e supervisão humana das decisões automatizadas (Silva, 2026).

2.5 Estado da arte e lacuna

A produção que fundamenta esta pesquisa pode ser organizada em três linhagens. A primeira é a doutrina civil-constitucional dos direitos da personalidade, que estabelece a autonomia do direito à imagem, sua dupla dimensão moral e patrimonial e a primazia da tutela preventiva (Schreiber, 2014; Teffé, 2017). A segunda é a teoria da responsabilidade civil, que exige analisar risco, causalidade e imputação a partir da atividade concretamente desempenhada (Tepedino; Terra; Guedes, 2026). A terceira, mais recente, é formada por artigos que tratam especificamente da inteligência artificial e do direito à imagem, e que registram o aumento da verossimilhança das manipulações, a dificuldade de identificação dos responsáveis e a ausência de parâmetros jurisprudenciais estabilizados sobre imputação, nexo causal e quantificação dos danos (Ebling; Ribeiro, 2024; Alves; Sampaio Neto, 2025; Amaral et al., 2026).

A essas três linhagens nacionais soma-se a literatura internacional que inaugurou o debate jurídico sobre o tema. Chesney e Citron (2019) mostram que os danos dos deepfakes não se limitam à vítima individual, cuja imagem pode ser explorada ou sabotada, e alcançam a confiança coletiva na prova audiovisual. Os autores descrevem o que denominam dividendo do mentiroso (liar’s dividend): quanto mais o público sabe que imagens e vozes podem ser fabricadas, mais fácil se torna, para quem foi flagrado em registro autêntico, alegar que se trata de falsificação. Examinam também os limites da responsabilidade civil como resposta, em razão da dificuldade de identificar os autores e da imunidade que o direito norte-

americano confere às plataformas por conteúdo de terceiros.

Os estudos recentes revelam, assim, que a dificuldade jurídica se deslocou da simples existência de tutela para a definição de quem deve responder. Alves e Sampaio Neto (2025) apontam problemas de identificação dos responsáveis, enquanto Amaral et al. (2026) concentram-se no ilícito lucrativo e registram a escassez de precedentes. Nenhuma dessas linhagens, porém, examina separadamente cada posição da cadeia tecnológica à luz do regime de plataformas redefinido pelo Supremo Tribunal Federal e regulamentado em 2026. A contribuição deste artigo é oferecer um conjunto de critérios de imputação aplicável a cada agente — criador, divulgador, plataforma e fornecedor —, sistematizado em matriz que permite distinguir quando cada um responde e quando não responde.

3 MATERIAL E MÉTODOS

A pesquisa utiliza o método dedutivo, partindo das normas gerais de proteção dos direitos da personalidade e dos pressupostos da responsabilidade civil para examinar sua aplicação às violações do direito à imagem produzidas por deepfakes. A abordagem é qualitativa e a investigação possui natureza jurídico-dogmática, pois interpreta normas, doutrina e decisões judiciais para identificar critérios de imputação compatíveis com o ordenamento brasileiro.

Quanto aos procedimentos, foram empregadas pesquisa bibliográfica, documental e jurisprudencial. A pesquisa bibliográfica reúne doutrina de direitos da personalidade e responsabilidade civil, com destaque para Schreiber (2014), Tepedino, Terra e Guedes (2026) e Teffé (2017), além de artigos científicos diretamente relacionados à inteligência artificial e ao direito à imagem, especialmente Ebling e Ribeiro (2024), Alves e Sampaio Neto (2025) e Amaral et al. (2026), e da literatura internacional sobre os aspectos técnicos e jurídicos dos deepfakes (Westerlund, 2019; Chesney; Citron, 2019).

A pesquisa documental abrange a Constituição Federal, o Código Civil, o Código de Defesa do Consumidor, o Marco Civil da Internet, a Lei Geral de Proteção de Dados, os Decretos nº

12.975/2026 e nº 12.976/2026 e o Projeto de Lei nº 2.338/2023. O exame das plataformas considera a tese vinculante dos Temas 987 e 533 do Supremo Tribunal Federal, aperfeiçoada no julgamento dos embargos de declaração concluído em 17 de junho de 2026, bem como a modulação dos efeitos e o debate de constitucionalidade instaurado na ADI 7.981. O recorte normativo e jurisprudencial foi atualizado até 12 de novembro de 2026 (Brasil, 2026a, 2026b, 2026c, 2026d, 2026e).

Na jurisprudência, foram utilizados diretamente o REsp 267.529/RJ e a Súmula 403 do Superior Tribunal de Justiça como parâmetros de proteção do direito à imagem. Quanto à Apelação Cível nº 5002715-53.2022.8.13.0443, do Tribunal de Justiça de Minas Gerais, e ao Processo nº 1029948-82.2024.8.26.0016, do Tribunal de Justiça de São Paulo, a análise é de segunda mão, com base na descrição feita por Amaral et al. (2026), razão pela qual esses julgados são identificados no corpo do texto com uso de apud e não integram autonomamente a lista de referências.

O material foi lido segundo três eixos, que correspondem às perguntas que a responsabilidade civil formula a cada agente: (i) a conduta, isto é, o que cada participante fez ou deixou de fazer na cadeia de criação, circulação e fornecimento da tecnologia; (ii) o dever, ou seja, qual norma atribui a esse participante um dever de agir, de prevenir ou de remover; e (iii) o regime, isto é, se a responsabilidade é subjetiva, objetiva, presumida ou solidária, e em quais condições. Os oito critérios apresentados na seção 4.7 foram formulados a partir desses eixos, tendo cada modalidade de deepfake identificada na seção 2.4 servido como caso de teste, e não resultam de um cruzamento formal entre eles.

O recorte material concentra-se na responsabilidade civil decorrente da violação do direito à imagem de pessoas vivas. Questões penais, eleitorais, autorais e relativas à reconstrução digital de pessoas falecidas aparecem apenas quando necessárias à compreensão do fenômeno. Também não se pretende avaliar tecnicamente a qualidade dos modelos de inteligência artificial, mas somente os aspectos tecnológicos indispensáveis à aplicação dos

institutos jurídicos.

Como limitação metodológica, reconhece-se a reduzida quantidade de precedentes brasileiros especificamente voltados a deepfakes e direito à imagem. Essa escassez impede afirmar a existência de jurisprudência consolidada para todas as modalidades de dano. Por isso, a pesquisa combina os precedentes disponíveis com a legislação vigente, a doutrina e a tese vinculante do STF aplicável às plataformas digitais.

4 RESULTADOS E DISCUSSÃO

4.1 Pressupostos da responsabilidade civil aplicáveis aos deepfakes

A utilização de tecnologia nova não afasta os elementos gerais da responsabilidade civil. Os arts. 186 e 187 do Código Civil tratam do ato ilícito e do abuso de direito, enquanto o art. 927 estabelece o dever de reparar. A responsabilidade subjetiva exige conduta, dano, nexo causal e culpa ou dolo; a objetiva dispensa a demonstração de culpa quando houver fundamento legal ou quando a atividade normalmente desenvolvida implicar risco especial para direitos de terceiros, sem afastar a necessidade de dano e nexo causal (Brasil, 2002).

Tepedino, Terra e Guedes (2026) destacam que risco, causalidade e imputação devem ser analisados a partir da atividade concretamente desempenhada, e não de forma abstrata. Essa cautela é especialmente importante em sistemas de inteligência artificial, nos quais diferentes agentes participam de etapas distintas. O simples fato de um conteúdo ter sido produzido com IA não transforma automaticamente todos os envolvidos em responsáveis objetivos.

O dano pode assumir natureza extrapatrimonial e patrimonial. A falsa representação pode atingir imagem, honra, intimidade, identidade e reputação, ao mesmo tempo em que pode gerar perda financeira ou exploração econômica indevida. A intensidade da lesão depende do conteúdo, da finalidade, do alcance, da duração da circulação e da credibilidade atribuída à

falsificação.

O nexo causal é um dos pontos mais sensíveis. Criador, divulgador, plataforma e fornecedor do sistema realizam condutas diferentes. A mera presença na cadeia tecnológica não basta para impor responsabilidade; é necessário verificar de que modo a ação ou omissão de cada agente contribuiu juridicamente para o dano e se havia dever de agir naquela situação.

A tutela dos direitos da personalidade também possui dimensão preventiva. O art. 12 do Código Civil admite providências para fazer cessar ameaça ou lesão, e o art. 944 determina que a indenização se mede pela extensão do dano. Em deepfakes de rápida circulação, remoção, preservação de prova, identificação do autor e impedimento de republicações podem ser tão relevantes quanto a reparação pecuniária (Brasil, 2002).

4.2 Responsabilidade do criador e do usuário que produz a manipulação

O criador humano do deepfake ocupa, em regra, a posição causal mais próxima da manipulação. Quando seleciona a imagem de terceiro, fornece comandos, realiza ajustes e produz deliberadamente conteúdo falso, sua atuação constitui conduta própria. Se o resultado viola direito da personalidade e causa dano, os pressupostos da responsabilidade civil podem ser examinados diretamente a partir dessa participação.

A automação não elimina a decisão humana de iniciar ou direcionar a produção. O contexto, porém, permanece relevante. Uma geração consentida para finalidade artística, educacional ou humorística claramente identificada não pode ser equiparada, sem análise, à falsificação destinada a enganar, explorar economicamente ou expor a vítima. Consentimento, finalidade e expectativa de autenticidade influenciam a ilicitude e a extensão do dano.

Quando a criação possui finalidade econômica, o proveito obtido ganha relevância adicional. Amaral et al. (2026) discutem o ilícito lucrativo em hipóteses nas quais a utilização indevida da imagem gera vantagem econômica ao ofensor. Ainda que essa teoria não resolva

automaticamente todos os casos, o ganho obtido pode influenciar a avaliação da conduta e da resposta reparatória.

O primeiro critério de distribuição da responsabilidade é, assim, a participação direta na criação. Quanto maior o domínio do agente sobre a escolha da vítima, a geração, a edição e a finalidade da publicação, mais nítida tende a ser sua contribuição causal. Isso não impede eventual coautoria de terceiros, mas evita diluir o autor direto em uma cadeia tecnológica abstrata.

4.3 Responsabilidade do divulgador e de quem amplia a circulação

Criação e divulgação constituem condutas distintas. Um conteúdo pode ser produzido por uma pessoa e divulgado por outra que não participou da etapa de geração. A responsabilidade do divulgador deve ser examinada conforme conhecimento da falsidade, finalidade do compartilhamento e contribuição efetiva para a ampliação do dano.

O compartilhamento consciente de material sintético apresentado como verdadeiro pode constituir nova agressão ao direito à imagem, sobretudo quando aumenta o alcance da falsificação ou reforça sua credibilidade. A situação é diferente quando o conteúdo é reproduzido para denúncia, pesquisa, crítica, verificação ou esclarecimento, com indicação clara de que se trata de montagem. A finalidade comunicativa e o contexto precisam ser considerados.

Também importa o momento em que o divulgador toma conhecimento da natureza manipulada do conteúdo. A continuidade da disseminação após ciência inequívoca da falsidade e da lesão pode aumentar a reprovabilidade da conduta. Em sentido contrário, a ausência de sinais objetivos de manipulação pode exigir exame mais rigoroso da culpa, especialmente diante de conteúdos tecnicamente sofisticados.

O critério central é a contribuição causal para a ampliação do dano, combinada ao

conhecimento e ao contexto. A responsabilização não deve ser presumida apenas porque alguém recebeu ou compartilhou um arquivo, mas pode surgir quando sua atuação contribui conscientemente para a permanência ou expansão da violação.

4.4 Responsabilidade das plataformas digitais: regimes diferenciados após os Temas 987 e 533 do STF

A responsabilidade das plataformas exige tratamento próprio, pois o provedor não se confunde com o usuário que cria o conteúdo. O Marco Civil da Internet estruturou originalmente dois regimes principais: o art. 19 condicionou, como regra, a responsabilidade por conteúdo de terceiros ao descumprimento de ordem judicial específica, enquanto o art. 21 estabeleceu notificação extrajudicial para divulgação não autorizada de imagens, vídeos ou outros materiais contendo cenas de nudez ou atos sexuais de caráter privado (Brasil, 2014).

Esse panorama foi profundamente alterado pelo Supremo Tribunal Federal. A mudança merece ser lida em perspectiva comparada: a imunidade das plataformas por conteúdo de terceiros, que Chesney e Citron (2019) apontam como um dos principais obstáculos à reparação das vítimas de deepfakes no direito norte-americano, foi precisamente o modelo do qual o STF se afastou ao reconhecer a inconstitucionalidade parcial e progressiva do art. 19. Após o julgamento de mérito de 2025, os embargos de declaração dos Temas 987 e 533 foram concluídos em 17 de junho de 2026, com ajuste da tese. O STF manteve a inconstitucionalidade parcial e progressiva do art. 19 e fixou regimes diferenciados de responsabilidade. Os efeitos foram modulados ex nunc a partir de 5 de agosto de 2025, data da publicação da ata do julgamento de mérito, ressalvados atos continuados ou permanentes e decisões transitadas em julgado (Brasil, 2026c, 2026d).

No regime geral, para crimes e atos ilícitos que não sejam contra a honra, a tese estendeu a lógica do art. 21: após notificação, o provedor que não adota providências pode responder

solidariamente pelos danos gerados por terceiros, salvo quando, após diligência qualificada, subsista dúvida razoável sobre a ilicitude. Para crimes e ilícitos civis contra a honra, permanece aplicável o art. 19, de modo que a responsabilização civil continua vinculada à ordem judicial específica, sem prejuízo da possibilidade de remoção voluntária após notificação extrajudicial (Brasil, 2026c, 2026d).

Outra hipótese autônoma diz respeito a anúncios, impulsionamentos pagos e mecanismos artificiais de disseminação inorgânica. Nos embargos, o STF substituiu a antiga expressão "presunção de responsabilidade" por presunção relativa de culpa. Nesses casos, a responsabilização pode ocorrer independentemente de notificação, mas o provedor pode afastá-la se demonstrar atuação diligente e em tempo razoável para tornar o conteúdo indisponível (Brasil, 2026c, 2026d).

Há ainda o dever de cuidado relacionado à circulação massiva de conteúdos ilícitos graves. A responsabilidade pressupõe falha sistêmica, e não a mera existência isolada de conteúdo. O rol inclui, entre outras hipóteses, crimes contra mulheres, crimes sexuais contra vulneráveis, pornografia infantil, terrorismo, tráfico de pessoas e determinadas condutas antidemocráticas. Essa distinção é relevante para deepfakes porque certas manipulações podem se enquadrar em categorias de elevada gravidade, enquanto outras permanecem submetidas aos regimes gerais (Brasil, 2026c, 2026d).

O Decreto nº 12.975/2026, publicado em 21 de maio de 2026 e em vigor sessenta dias depois, regulamentou deveres estruturais, canais de denúncia, tratamento de notificações, falha sistêmica, anúncios e impulsionamentos. O texto do art. 16-L emprega a expressão "presume-se a responsabilidade", mas a tese vinculante posteriormente ajustada pelo STF passou a falar em presunção relativa de culpa; o parágrafo único do mesmo artigo, aliás, já excluía a responsabilidade do provedor que comprovasse atuação diligente e em tempo razoável (Brasil, 2026a, art. 16-L). A constitucionalidade do decreto é objeto da ADI 7.981, na qual se questiona, entre outros pontos, eventual inovação normativa além dos limites

regulamentares e se pede interpretação compatível com a tese do STF. Tramitam também projetos de decreto legislativo destinados a sustar seus efeitos, como os PDLs nº 396/2026 e nº 407/2026, na Câmara dos Deputados, e o PDL nº 398/2026, no Senado Federal, o que demonstra que sua legalidade permanece em debate (Brasil, 2026a, 2026e, 2026f, 2026g, 2026h).

O Decreto nº 12.976/2026, também publicado em 21 de maio e em vigor sessenta dias depois, adota a definição de conteúdo íntimo referida na introdução, que abrange o material sintético. Seu art. 4º prevê a responsabilização dos provedores em caso de falha sistêmica na indisponibilização de conteúdos que configurem crimes ou atos ilícitos praticados contra mulheres em razão da condição do sexo feminino. O art. 7º determina a indisponibilização do conteúdo íntimo em até duas horas após a notificação (§ 1º), em toda a aplicação, com marcação digital para que o reenvio seja automaticamente bloqueado (§ 4º), e os arts. 9º e 10 vedam aos provedores a geração ou modificação de conteúdo íntimo de terceiro por IA e exigem salvaguardas técnicas e procedimentais para identificar e bloquear solicitações dessa natureza. Essas regras dialogam diretamente com deepfakes sexuais e reforçam a necessidade de distinguir a modalidade do conteúdo antes de definir o dever da plataforma ou do fornecedor (Brasil, 2026b).

Quadro 1 – Regimes aplicáveis às principais modalidades de deepfake

Modalidade	Regime predominante	Consequência principal
Conteúdo íntimo ou sexual não consentido	Art. 21 do MCI; para mulheres, Decreto nº 12.976/2026	Notificação extrajudicial; no Decreto nº 12.976/2026, indisponibilização em até duas horas, bloqueio de reenvio e salvaguardas contra geração por IA.
Ataque à honra / falsa fala de conteúdo difamatório, calunioso ou injurioso	Art. 19 do MCI, conforme tese final dos Temas 987 e 533	Responsabilização civil da plataforma depende de ordem judicial específica, sem impedir remoção voluntária após notificação.

Modalidade	Regime predominante	Consequência principal
Fraude ou outro crime/ato ilícito não enquadrado como honra	Regime geral fixado pelo STF, por extensão da lógica do art. 21	Após notificação, pode haver responsabilidade solidária se o provedor não agir, salvo dúvida razoável após diligência qualificada.
Anúncio, impulsionamento pago ou disseminação inorgânica artificial	Temas 987 e 533; Decreto nº 12.975/2026, art. 16-L	Presunção relativa de culpa; não exige notificação prévia; admite afastamento mediante prova de atuação diligente e tempestiva.
Circulação massiva de conteúdos ilícitos graves	Dever de cuidado / falha sistêmica	Responsabilidade depende de falha sistêmica na prevenção ou remoção das categorias graves previstas na tese e na regulamentação.

Fonte: elaboração própria, com base em Brasil (2014, 2026a, 2026b, 2026c, 2026d).

O quadro evidencia que a expressão “responsabilidade da plataforma por deepfake” não descreve um regime único. A correta imputação depende, primeiro, da classificação jurídica do conteúdo e, depois, da conduta concreta do provedor. Essa diferenciação sustenta a tese de individualização adotada neste trabalho.

4.5 Responsabilidade do fornecedor do sistema de inteligência artificial

O fornecedor do sistema generativo ocupa posição diferente da plataforma que hospeda ou distribui o conteúdo. Em alguns serviços, geração e publicação estão integradas; em outros, o usuário produz o material em uma ferramenta e o transfere para rede distinta. A responsabilização exige, portanto, reconstruir a efetiva participação do fornecedor e os deveres jurídicos vinculados ao serviço prestado.

Na responsabilidade civil geral, fatores como previsibilidade do uso lesivo, grau de controle técnico, existência de salvaguardas razoáveis, resposta a denúncias, possibilidade de bloquear comandos manifestamente abusivos e proveito econômico ajudam a identificar

eventual conduta ou omissão relevante. A teoria do risco do art. 927, parágrafo único, não deve ser aplicada apenas porque a tecnologia é sofisticada; é necessário verificar se a atividade normalmente desenvolvida, em sua configuração concreta, cria risco juridicamente relevante para direitos de terceiros (Brasil, 2002; Tepedino; Terra; Guedes, 2026).

A análise se fortalece quando houver relação de consumo. O art. 14 do Código de Defesa do Consumidor estabelece responsabilidade objetiva do fornecedor por defeito do serviço, observadas as excludentes legais, e o art. 17 equipara a consumidor todas as vítimas do evento. Assim, uma pessoa atingida por defeito de segurança ou funcionamento do serviço pode, conforme o caso concreto, ser protegida como consumidora por equiparação mesmo sem relação contratual direta com o fornecedor. Isso não significa responsabilidade automática por qualquer deepfake, mas impede tratar o CDC como hipótese meramente marginal quando o dano decorre do próprio serviço (Brasil, 1990).

O Decreto nº 12.976/2026 tornou esse debate ainda mais concreto no campo de conteúdo íntimo envolvendo mulheres. A norma veda a geração e modificação de conteúdo íntimo de terceiro por IA e exige salvaguardas técnicas e procedimentais destinadas a identificar e bloquear solicitações proibidas, implementadas de forma escalonada e proporcional ao volume de acessos e ao nível de risco da aplicação (Brasil, 2026b, art. 10, parágrafo único). Para serviços abrangidos, a ausência ou ineficácia dessas medidas pode integrar a análise sobre defeito, diligência, prevenção e nexo causal, sem afastar a necessidade de examinar as circunstâncias do caso.

No plano legislativo, o Projeto de Lei nº 2.338/2023 segue em tramitação na Câmara dos Deputados, como se viu na seção 2.4. Por não estar aprovado, não pode ser tratado como fundamento jurídico vigente, mas funciona como indicador de que o Congresso ainda discute parâmetros gerais de governança, risco e responsabilização aplicáveis a sistemas de IA (Brasil, 2023).

A conclusão, portanto, é que o fornecedor deve ser individualizado. Sua posição não se confunde com a do usuário que deliberadamente produz a falsificação, mas também não pode ser considerada neutra em todas as hipóteses, especialmente quando o dano se relaciona a defeito do serviço, ausência de salvaguardas exigíveis ou descumprimento de dever normativo específico.

4.6 Precedentes relacionados ao direito à imagem e a deepfakes

A jurisprudência do Superior Tribunal de Justiça fornece base consolidada para a proteção da imagem. O REsp 267.529/RJ e a Súmula 403, examinados na seção 2.1, demonstram que a tutela da imagem antecede o fenômeno dos deepfakes e pode ser aplicada às novas formas de apropriação sintética, desde que respeitadas as particularidades do caso (Brasil, 2000, 2009).

Quanto a precedentes especificamente relacionados a deepfakes, a literatura ainda registra número reduzido de casos. Conforme descrito por Amaral et al. (2026), a Apelação Cível nº 5002715-53.2022.8.13.0443, do Tribunal de Justiça de Minas Gerais, examinou prejuízos vinculados a empréstimo obtido mediante emprego de deepfake, demonstrando que a manipulação sintética pode integrar fraude patrimonial e exigir reconstrução da cadeia causal (Minas Gerais, 2024 apud Amaral et al., 2026).

Também conforme Amaral et al. (2026), o Processo nº 1029948-82.2024.8.26.0016, do Tribunal de Justiça de São Paulo, envolveu conteúdo de nudez produzido por deepfake e discussão sobre a conduta da plataforma após notificação. O caso evidencia que a responsabilidade do intermediário não decorre de ter criado a falsificação, mas de sua atuação ou omissão após adquirir ciência do conteúdo apontado como ilícito (São Paulo, 2025 apud Amaral et al., 2026).

Esses julgados são utilizados com cautela porque, neste trabalho, não foram consultados diretamente em inteiro teor; a descrição dos fatos e fundamentos decorre da fonte

acadêmica secundária. Ainda assim, eles ilustram duas estruturas distintas — fraude patrimonial e conteúdo íntimo sintético — e reforçam a necessidade de não tratar todos os deepfakes sob um único modelo de imputação.

4.7 Critérios para distribuir a responsabilidade entre os diferentes agentes

A partir da legislação, da doutrina, da jurisprudência e dos regimes específicos de plataformas, é possível sistematizar critérios para responder ao problema de pesquisa. O primeiro é a participação causal: deve-se identificar quem criou, ordenou, editou, publicou, impulsionou, manteve ou deixou de remover o conteúdo e de que modo essa conduta se relaciona com o dano. Quanto mais direta a intervenção, mais evidente tende a ser a conexão causal, embora a participação de um agente não exclua a de outros.

O segundo critério é o conhecimento da ilicitude. Para o autor que deliberadamente fabrica conteúdo falso, essa ciência normalmente decorre da própria criação. Para divulgadores, plataformas e fornecedores, o conhecimento pode resultar do contexto, de notificações, de sinais objetivos ou de deveres específicos. A dúvida razoável sobre a ilicitude assume especial importância no regime das plataformas após os Temas 987 e 533 (Brasil, 2026c, 2026d).

O terceiro critério é a existência de dever jurídico de agir. A omissão somente produz responsabilidade quando o ordenamento atribui ao agente um dever de comportamento. Esse dever pode decorrer do Código Civil, do CDC, do Marco Civil, de normas de proteção de dados, dos Decretos nº 12.975/2026 e nº 12.976/2026 ou da própria tese vinculante do STF, conforme a posição do agente e a natureza do conteúdo.

O quarto critério é a capacidade efetiva de prevenção ou mitigação. Deve-se avaliar se o agente possui meios técnicos e jurídicos para reduzir o risco ou interromper a lesão. Não se pode impor obrigação impossível, mas a existência de mecanismos de bloqueio, denúncia, rastreabilidade, autenticação, prevenção de reenvio ou filtragem pode ser relevante quando

o risco é previsível e a resposta está sob controle do agente.

O quinto critério é o proveito econômico. Publicidade, impulsionamento, monetização e exploração comercial da imagem podem influenciar a avaliação da conduta e da reparação. A Súmula 403 do STJ demonstra que o uso econômico não autorizado possui tutela consolidada, e Amaral et al. (2026) acrescentam a discussão sobre ilícito lucrativo como instrumento para evitar que a violação permaneça economicamente vantajosa.

O sexto critério é a natureza e a extensão do dano. Conteúdo íntimo, fraude patrimonial, publicidade enganosa e falsa atribuição de comportamento possuem estruturas diferentes. Gravidade, alcance, duração da circulação, possibilidade de replicação e efeitos profissionais ou pessoais devem ser considerados tanto para a imputação quanto para a quantificação da reparação.

O sétimo critério é o contexto de consentimento e autenticidade. A vítima pode autorizar determinado uso de sua imagem sem consentir com transformação ampla ou finalidade diversa. A disponibilização de fotografia na internet não equivale a autorização genérica para criar representações sintéticas. Conteúdo claramente identificado como ficcional ou satírico também apresenta contexto distinto daquele construído para ser confundido com registro verdadeiro.

O oitavo critério é a possibilidade de identificação e prova. Registros, dados de acesso, histórico de notificações, informações sobre impulsionamento e elementos técnicos podem ser essenciais para reconstruir a cadeia de criação e circulação. O Marco Civil já impõe aos provedores de aplicações constituídos como pessoa jurídica a guarda dos registros de acesso pelo prazo de seis meses e permite à parte interessada requerer ao juiz o fornecimento desses registros para formar prova (Brasil, 2014, arts. 15 e 22), e o Decreto nº 12.975/2026 obriga os provedores que oferecem impulsionamento pago a manter, por um ano após o encerramento da veiculação, as informações relativas a cada anúncio e a cada anunciante

(Brasil, 2026a, art. 16-M). A prova tem ainda uma dimensão inversa, que o dividendo do mentiroso descrito por Chesney e Citron (2019) torna visível: a mesma tecnologia que permite fabricar registros falsos permite ao ofensor desacreditar registros verdadeiros. A preservação de metadados, registros de criação e histórico de circulação serve, portanto, tanto à vítima do deepfake quanto à vítima de um registro autêntico que se pretende desqualificar como sintético. A opacidade tecnológica não deve justificar responsabilização indiscriminada, mas tampouco pode funcionar como mecanismo de impunidade quando determinado agente detém informações necessárias à apuração.

A rejeição de uma solidariedade indistinta não significa afastar toda solidariedade. O art. 942 do Código Civil determina, em seu caput, que, se a ofensa tiver mais de um autor, todos responderão solidariamente pela reparação, e seu parágrafo único estende a solidariedade aos coautores. Assim, se diferentes agentes forem juridicamente caracterizados como coautores ou participantes do mesmo ilícito, a solidariedade pode incidir. Além disso, a tese final do STF estabeleceu hipótese específica de responsabilidade solidária da plataforma no regime geral após notificação, ressalvada dúvida razoável. O que se rejeita é a ideia de que todos os integrantes da cadeia tecnológica respondem solidariamente apenas porque participaram, de algum modo, do ecossistema que tornou o deepfake possível (Brasil, 2002, 2026c, 2026d).

O Quadro 2 aplica os oito critérios às quatro posições da cadeia, indicando, para cada agente, a pergunta que o critério formula e o fundamento normativo que a orienta.

Quadro 2 – Aplicação dos critérios de imputação a cada agente da cadeia do deepfake

Critério	Criador	Divulgador	Plataforma	Fornecedor de IA
Participação causal	Seleciona a vítima, comanda e edita a geração (CC, arts. 186 e 927)	Amplia o alcance ou reforça a credibilidade	Hospeda, recomenda, impulsiona ou mantém	Fornece a ferramenta; integra ou não geração e publicação
Conhecimento da ilicitude	Decorre da própria criação	Ciência da falsidade e momento em que ela ocorre	Notificação, sinais objetivos, dúvida razoável (Temas 987 e 533)	Previsibilidade do uso lesivo; denúncias recebidas
Dever jurídico de agir	Dever geral de não lesar (CC, arts. 186 e 927)	Cessar a divulgação após ciência	MCI, arts. 19 e 21; Decretos nº 12.975 e nº 12.976/2026	Segurança do serviço (CDC, art. 14); Decreto nº 12.976/2026, arts. 9º e 10
Capacidade de prevenção ou mitigação	Plena: basta não produzir	Limitada: deixar de compartilhar	Remoção, bloqueio de reenvio, redução de alcance	Salvaguardas contra comandos abusivos
Proveito econômico	Ilícito lucrativo; Súmula 403 do STJ	Monetização do compartilhamento	Anúncios e impulsionamentos pagos (presunção relativa de culpa)	Remuneração pelo serviço de geração
Natureza e extensão do dano	Define o regime aplicável (Quadro 1) e a quantificação (CC, art. 944)	Idem, considerado o alcance acrescido	Idem, considerada a duração da circulação	Idem, considerado o defeito do serviço

Critério	Criador	Divulgador	Plataforma	Fornecedor de IA
Consentimento e autenticidade	Finalidade, contexto e sinalização de ficção ou sátira	Indicação de que se trata de montagem	Rotulagem e contexto da publicação	Consentimento exigível para conteúdo íntimo
Identificação e prova	Autor a identificar	Registros de compartilhamento	Guarda de registros (MCI, arts. 15 e 22; Decreto nº 12.975/2026, art. 16-M)	Registros de solicitações e de bloqueios

Fonte: elaboração própria, com base em Brasil (1990, 2002, 2009, 2014, 2026a, 2026b, 2026c, 2026d).

A leitura do quadro em colunas mostra o perfil de cada agente: o criador concentra a participação causal e o conhecimento; o divulgador responde pela ampliação consciente; a plataforma, pelo descumprimento de deveres que variam conforme o regime do conteúdo; e o fornecedor, pelo defeito do serviço ou pela ausência de salvaguardas exigíveis. A leitura em linhas mostra que nenhum critério isolado decide a imputação: é sua combinação que permite afirmar, em cada caso, quem responde e por quê.

4.8 Reparação, prevenção e efetividade da tutela

A reparação civil deve considerar a natureza específica do dano. O art. 944 do Código Civil determina que a indenização se mede pela extensão da lesão, enquanto a Súmula 403 do STJ afasta a necessidade de prova do prejuízo quando a imagem é utilizada sem autorização para fins econômicos ou comerciais. Em deepfakes, a quantificação pode considerar alcance, duração, gravidade, repercussão, finalidade econômica e possibilidade de replicação (Brasil, 2002, 2009).

Amaral et al. (2026) observam que ainda não há parâmetros jurisprudenciais estabilizados

para quantificar danos especificamente decorrentes de deepfakes. Quando a violação produz lucro, a resposta civil também pode considerar a necessidade de impedir que o ilícito permaneça economicamente vantajoso, sem transformar a indenização em sanção automática desvinculada dos institutos aplicáveis.

A efetividade depende igualmente de prevenção. Conteúdos sintéticos podem continuar circulando depois de reconhecida a falsidade, e a remoção tardia não recompõe integralmente a imagem, a privacidade ou a reputação atingidas. Por isso, tutela inibitória, indisponibilização célere, preservação de provas, identificação do responsável e mecanismos contra reenvio ganham importância. Os Decretos nº 12.975/2026 e nº 12.976/2026 reforçam essa dimensão preventiva ao disciplinar notificações, deveres estruturais e, no caso de conteúdo íntimo contra mulheres, prazos curtos de indisponibilização e salvaguardas para sistemas de IA (Brasil, 2026a, 2026b).

A tutela deve, contudo, preservar liberdade de expressão, devido processo e segurança jurídica. A divisão de regimes feita pelo STF, a possibilidade de dúvida razoável e a manutenção do art. 19 para ilícitos contra a honra demonstram que a proteção da imagem não autoriza remoção indiscriminada. O desafio está em construir respostas proporcionais ao tipo de conteúdo, ao grau de evidência da ilicitude e à posição de cada participante.

4.9 Objeções

A primeira objeção, em sua versão mais forte, sustenta que a individualização protege mal a vítima: diante de um autor anônimo e de uma cadeia tecnológica difusa, a solução mais protetiva seria a responsabilidade objetiva e solidária de todos os que lucram com o ecossistema. A objeção tem apelo, mas desconsidera dois pontos. O primeiro é que a vítima não fica sem devedor solvente: a plataforma responde solidariamente no regime geral após notificação, sob presunção relativa de culpa nos anúncios e impulsionamentos, e o fornecedor responde objetivamente quando há defeito do serviço e a vítima é equiparada a

consumidora (seções 4.4 e 4.5). O segundo é que a solidariedade indistinta não encontra fundamento no art. 942 do Código Civil, que a reserva aos autores e coautores da ofensa, e transformaria em garantes universais agentes que não tinham conhecimento, dever de agir nem capacidade de evitar o dano, deslocando o custo do ilícito para quem não o praticou.

A segunda objeção afirma que critérios fundados no conhecimento e no dever de agir seriam inaplicáveis na prática, porque o criador de deepfakes costuma ser anônimo e a reconstrução da cadeia, impossível. A dificuldade é real, e foi registrada tanto pela literatura nacional (Alves; Sampaio Neto, 2025) quanto pela internacional, que a aponta como um dos limites da responsabilidade civil diante dos deepfakes (Chesney; Citron, 2019). Ela não decorre, porém, da ausência de instrumentos, mas da sua subutilização. A guarda obrigatória de registros de acesso, a possibilidade de requisição judicial desses registros e o dever de manter as informações sobre anúncios e anunciantes (seção 4.7) existem justamente para permitir a identificação. O oitavo critério responde a essa objeção ao converter a posse dessas informações em fator de imputação: o agente que as detém e não as preserva contribui para a impunidade que a objeção descreve.

A terceira objeção vem do lado oposto e teme que a ampliação dos deveres de plataformas e fornecedores leve à remoção excessiva de conteúdos legítimos, como sátiras, paródias e críticas políticas que empregam manipulação de imagem. A resposta está no sétimo critério e no desenho do próprio regime. O contexto de consentimento e autenticidade separa o conteúdo claramente identificado como ficcional daquele construído para ser confundido com registro verdadeiro, e a manutenção do art. 19 para ilícitos contra a honra, ao lado da excludente de dúvida razoável, impede que a plataforma seja responsabilizada por não remover conteúdo cuja ilicitude não é evidente. A individualização, nesse sentido, protege também a liberdade de expressão.

5 CONSIDERAÇÕES FINAIS

O desenvolvimento da inteligência artificial generativa ampliou a capacidade de manipular imagem, voz e comportamento de pessoas reais, tornando possível criar conteúdos sintéticos com alto grau de verossimilhança. O problema jurídico não está apenas na falsificação, mas na definição de quem deve responder quando criação, divulgação, hospedagem, impulsionamento e fornecimento da tecnologia envolvem agentes distintos.

A pesquisa confirma a hipótese apresentada na introdução. O ordenamento jurídico brasileiro possui fundamentos aptos a proteger o direito à imagem e assegurar tutela preventiva e reparatória, mas a imputação não pode ser uniforme. A resposta ao problema de pesquisa pode ser organizada em oito critérios: participação causal; conhecimento da ilicitude; existência de dever jurídico de agir; capacidade de prevenção ou mitigação; proveito econômico; natureza e extensão do dano; contexto de consentimento e autenticidade; e possibilidade de identificação e prova.

O criador ou usuário que deliberadamente produz a manipulação tende a apresentar vínculo causal mais direto. O divulgador deve ser avaliado conforme conhecimento, finalidade e contribuição para a ampliação da lesão. O fornecedor de IA exige exame sobre desenho do serviço, controle técnico, previsibilidade do risco, salvaguardas e eventual incidência do Código de Defesa do Consumidor, especialmente dos arts. 14 e 17 quando houver defeito do serviço e vítima equiparada a consumidor. A posição das plataformas tornou-se mais complexa após o julgamento dos Temas 987 e 533 e os embargos concluídos em 2026: ilícitos contra a honra permanecem vinculados ao art. 19; crimes e outros atos ilícitos seguem regime geral de notificação, com responsabilidade solidária nas condições fixadas pela tese; anúncios, impulsionamentos pagos e disseminação inorgânica atraem presunção relativa de culpa; e conteúdos graves em circulação massiva são examinados sob a lógica da falha sistêmica. Para conteúdo íntimo, o art. 21 do Marco Civil e o Decreto nº 12.976/2026 oferecem tutela específica, inclusive para material manipulado por inteligência artificial (Brasil, 2014, 2026b, 2026c, 2026d).

A solidariedade também exige precisão. Não há fundamento para impor responsabilidade solidária automática a todos os participantes da cadeia apenas por sua presença no ecossistema tecnológico. Entretanto, o art. 942 do Código Civil admite solidariedade entre autores e coautores do ilícito, e a própria tese do STF prevê solidariedade da plataforma em determinadas hipóteses de omissão após notificação. A individualização, portanto, não exclui a solidariedade; ela define quando existem fundamentos concretos para aplicá-la.

A contribuição do trabalho está em duas sínteses. A primeira é a demonstração, no Quadro 1, de que a expressão "responsabilidade da plataforma por deepfake" abriga ao menos cinco regimes distintos, conforme a modalidade do conteúdo. A segunda é a matriz do Quadro 2, que aplica os oito critérios a cada posição da cadeia e permite, em cada caso, identificar quem responde, com qual fundamento e por que os demais não respondem.

Os limites da pesquisa também precisam ser reconhecidos. Os precedentes brasileiros diretamente voltados a deepfakes ainda são escassos, e os dois casos estaduais examinados foram conhecidos por fonte secundária, o que impede afirmar a existência de parâmetros estabilizados para nexo causal, quantificação do dano e distribuição da responsabilidade. Além disso, os Decretos nº 12.975/2026 e nº 12.976/2026 estão em vigor há poucos meses, e a constitucionalidade do primeiro é discutida na ADI 7.981, de modo que parte das conclusões sobre deveres das plataformas depende do desfecho desse julgamento. As conclusões sobre conteúdo íntimo, apoiadas no Decreto nº 12.976/2026, não são objeto dessa ação, embora ambos os decretos sejam alvo de projetos de decreto legislativo.

Conclui-se que a proteção do direito à imagem diante dos deepfakes não depende, no presente momento, da criação de um regime civil completamente apartado dos institutos existentes. O desafio está em aplicar de modo coerente as normas vigentes, os regimes específicos das plataformas e os deveres de segurança dos fornecedores, sem perder de vista liberdade de expressão, devido processo e inovação tecnológica. Pesquisas futuras deverão acompanhar a aplicação prática dos Decretos nº 12.975/2026 e nº 12.976/2026, o

julgamento da ADI 7.981, a evolução do PL nº 2.338/2023 e a formação de jurisprudência específica sobre deepfakes, inclusive por meio de levantamento sistemático das decisões dos tribunais estaduais, como o Tribunal de Justiça de Rondônia, que permita testar a matriz aqui proposta em casos concretos.

Conflito de interesses: os autores declaram não haver conflito de interesses.

Financiamento: a pesquisa não recebeu financiamento de agências de fomento públicas, comerciais ou sem fins lucrativos.

REFERÊNCIAS

ALVES, Adriano Mateus; SAMPAIO NETO, Walter Francisco. Direito de imagem, o combate e proteção da pessoa que tem sua imagem lesada por meio do sistema de inteligência artificial em dias atuais. **Revista Linhas Jurídicas**, Votuporanga, v. 12, n. 1, p. 171-200, jan./jun. 2025. Disponível em: https://periodicos.unifev.edu.br/linhasjuridicas/article/view/1985. Acesso em: 12 nov. 2026.

AMARAL, Ana Cláudia Zuin Mattos do; CASTRO, Camila Vieira; COSTA, Izabella Affonso; TEBAR, Natávia Boigues Corbalan. Deepfakes, o uso do direito de imagem e a responsabilidade civil pelo ilícito lucrativo. **Revista de Direito Contemporâneo UNIDEP**, Pato Branco, v. 5, n. 1, 2026. DOI: 10.68431/rdc.v5i1.434. Disponível em: https://periodicos.unidep.edu.br/rdc-u/article/view/434. Acesso em: 12 nov. 2026.

BRASIL. [Constituição (1988)]. **Constituição da República Federativa do Brasil de 1988**. Brasília, DF: Presidência da República, 1988. Disponível em: https://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 12 nov. 2026.

BRASIL. Lei nº 8.078, de 11 de setembro de 1990. Código de Defesa do Consumidor. Brasília, DF: Presidência da República, 1990. Disponível em:

https://www.planalto.gov.br/ccivil_03/leis/l8078compilado.htm. Acesso em: 12 nov. 2026.

BRASIL. Superior Tribunal de Justiça. Recurso Especial nº 267.529/RJ. Relator: Ministro Sálvio de Figueiredo Teixeira. Quarta Turma. Julgado em 3 out. 2000. **Diário da Justiça**, 18 dez. 2000, p. 208. Disponível em: https://scon.stj.jus.br/SCON/pesquisar.jsp?b=ACOR&fr=veja&i=1&livre=RESP+INPATH+%28CLAS%29+AND+267529+INPATH+%28NUM%29. Acesso em: 12 nov. 2026.

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: 12 nov. 2026.

BRASIL. Superior Tribunal de Justiça. Súmula 403. Independe de prova do prejuízo a indenização pela publicação não autorizada de imagem de pessoa com fins econômicos ou comerciais. Segunda Seção, julgado em 28 out. 2009, **DJe**, 24 nov. 2009. Disponível em: https://scon.stj.jus.br/SCON/sumstj/doc.jsp?b=SUMU&i=1&l=10&livre=%22403%22+INPATH%28NUM%29&operador=AND&ordenacao=-%40NUM&p=false. Acesso em: 12 nov. 2026.

BRASIL. Lei nº 12.965, de 23 de abril de 2014. Marco Civil da Internet. Brasília, DF: Presidência da República, 2014. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l12965.htm. Acesso em: 12 nov. 2026.

BRASIL. Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais. Brasília, DF: Presidência da República, 2018. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm. Acesso em: 12 nov. 2026.

BRASIL. Câmara dos Deputados. Projeto de Lei nº 2.338, de 2023. Dispõe sobre o

desenvolvimento, o fomento e o uso ético e responsável da inteligência artificial com base na centralidade da pessoa humana. Situação: aguardando parecer do relator na Comissão Especial sobre Inteligência Artificial. Brasília, DF: Câmara dos Deputados, 2023. Disponível em: https://www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=2487262. Acesso em: 12 nov. 2026.

BRASIL. Presidência da República. Decreto nº 12.975, de 20 de maio de 2026. Altera o Decreto nº 8.771, de 11 de maio de 2016, que regulamenta a Lei nº 12.965, de 23 de abril de 2014. **Diário Oficial da União**, Brasília, DF, 21 maio 2026a. Disponível em: https://www2.camara.leg.br/legin/fed/decret/2026/decreto-12975-20-maio-2026-799136-publicacaooriginal-179396-pe.html. Acesso em: 12 nov. 2026.

BRASIL. Presidência da República. Decreto nº 12.976, de 20 de maio de 2026. Estabelece diretrizes para a proteção de mulheres na internet e para o enfrentamento da violência contra mulheres em ambiente digital. **Diário Oficial da União**, Brasília, DF, 21 maio 2026b. Disponível em: https://www2.camara.leg.br/legin/fed/decret/2026/decreto-12976-20-maio-2026-799137-publicacaooriginal-179397-pe.html. Acesso em: 12 nov. 2026.

BRASIL. Supremo Tribunal Federal. Recurso Extraordinário nº 1.037.396/SP. Tema 987 da repercussão geral. Embargos de declaração. Tese aperfeiçoada em julgamento concluído em 17 jun. 2026. Brasília, DF: STF, 2026c. Disponível em: https://portal.stf.jus.br/jurisprudenciaRepercussao/verAndamentoProcesso.asp?incidente=5160549&numeroTema=987. Acesso em: 12 nov. 2026.

BRASIL. Supremo Tribunal Federal. Recurso Extraordinário nº 1.057.258/MG. Tema 533 da repercussão geral. Julgamento conjunto com o Tema 987; tese aperfeiçoada em embargos de declaração concluídos em 17 jun. 2026. Brasília, DF: STF, 2026d. Disponível em: https://portal.stf.jus.br/jurisprudenciaRepercussao/tema.asp?num=533. Acesso em: 12 nov.

2026.

BRASIL. Supremo Tribunal Federal. Ação Direta de Inconstitucionalidade nº 7.981/DF. Relator: Ministro Alexandre de Moraes. Questiona dispositivos do Decreto nº 12.975/2026. Brasília, DF: STF, 2026e. Disponível em: https://portal.stf.jus.br/processos/detalhe.asp?incidente=7631730. Acesso em: 12 nov. 2026.

BRASIL. Câmara dos Deputados. Projeto de Decreto Legislativo nº 396, de 2026. Susta os efeitos dos Decretos nº 12.975 e nº 12.976, de 20 de maio de 2026. Brasília, DF: Câmara dos Deputados, 2026f. Disponível em: https://www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=2626867. Acesso em: 15 nov. 2026.

BRASIL. Câmara dos Deputados. Projeto de Decreto Legislativo nº 407, de 2026. Susta os efeitos do Decreto nº 12.975, de 20 de maio de 2026. Brasília, DF: Câmara dos Deputados, 2026g. Disponível em: https://www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=2626855. Acesso em: 15 nov. 2026.

BRASIL. Senado Federal. Projeto de Decreto Legislativo nº 398, de 2026. Sustação dos efeitos do Decreto nº 12.975, de 20 de maio de 2026. Brasília, DF: Senado Federal, 2026h. Disponível em: https://www25.senado.leg.br/web/atividade/materias/-/materia/174259. Acesso em: 15 nov. 2026.

CHESNEY, Robert; CITRON, Danielle Keats. Deep fakes: a looming challenge for privacy, democracy, and national security. **California Law Review**, Berkeley, v. 107, n. 6, p. 1753-1820, 2019. DOI: 10.15779/Z38RV0D15J.

EBLING, Maurício; RIBEIRO, Luiz Henrique de Souza. O impacto no direito à imagem com os avanços da inteligência artificial. **Tecnologias em Projeção**, v. 15, n. 1, e1524TE02, 2024.

Disponível em: https://revistatecnologias.com/index.php/Projecao4/article/view/2432. Acesso em: 12 nov. 2026.

SCHREIBER, Anderson. **Direitos da personalidade**. 3. ed. São Paulo: Atlas, 2014.

SILVA, Renato Luiz Gomez da. Inteligência artificial pública e soberana no Poder Judiciário brasileiro: arquitetura institucional e epistemológica para a mitigação do racismo algorítmico. **Revista da Emeron**, Porto Velho, v. 3, p. 143-173, 2026. DOI: 10.62009/Emeron.2764.9679.v3.2026.529.p143-173.

TEFFÉ, Chiara Antonia Spadaccini de. Considerações sobre a proteção do direito à imagem na internet. **Revista de Informação Legislativa**, Brasília, v. 54, n. 213, p. 173-198, jan./mar. 2017. Disponível em: https://www12.senado.leg.br/ril/edicoes/54/213/ril_v54_n213_p173. Acesso em: 12 nov. 2026.

TEPEDINO, Gustavo; TERRA, Aline de Miranda Valverde; GUEDES, Gisela Sampaio da Cruz. **Fundamentos do direito civil**: responsabilidade civil. 7. ed. Rio de Janeiro: Forense, 2026. v. 4.

WESTERLUND, Mika. The emergence of deepfake technology: a review. **Technology Innovation Management Review**, Ottawa, v. 9, n. 11, p. 39-52, 2019. DOI: 10.22215/timreview/1282.

---

[1] Acadêmico do Curso de Direito do Centro Universitário São Lucas – Porto Velho.

[2] Orientador de Trabalho de Conclusão de Curso. Docente do Curso de Direito do Centro Universitário São Lucas – Afya Porto Velho, Doutorando em Letras (PPGL/UNIR), Mestre em Estudos Literários (UNIR), Especialista em Revisão de Textos (Unyleya) e Analista Judiciário – Revisor Redacional do Tribunal de Justiça do Estado de Rondônia (TJRO), atuando na Divisão

de Estudos e Diagnósticos de Aprimoramento do 1º Grau (Diap1G).