
FAKE NEWS COMO PROBLEMA JURÍDICO E SEU MÉTODO DE VERIFICAÇÃO JUDICIAL

*FAKE NEWS AS A LEGAL ISSUE AND ITS
METHOD OF JUDICIAL VERIFICATION*

Diogo Rais¹

¹ Doutor e mestre em Direito pela Pontifícia Universidade Católica de São Paulo (PUC-SP). Especialista em Direito Empresarial pela Universidade São Francisco (USF). Bacharel em Direito pela USF. Professor de Direito Eleitoral e Direito Digital da graduação, do mestrado e doutorado da Universidade Presbiteriana Mackenzie.

SUMÁRIO: Introdução. 1. O ponto de partida: o que o artigo 323 do Código Eleitoral tentou resolver e por que não resolveu. 2. A pergunta que muda tudo: o direito pune a mentira ou o dano? 3. *Fake news* como conteúdo enganoso com potencial lesivo. 4. Os dois passos do método de verificação judicial e análise de casos concretos. 5. O impacto da inteligência artificial generativa e os bens jurídicos envolvidos. Conclusão. Referências.

RESUMO: O artigo propõe a compreensão jurídica de *fake news* como conteúdo enganoso com potencial lesivo, cujo ponto de partida é a oposição ao conceito de “sabidamente inverídico” derivado do artigo 323 do Código Eleitoral, conceito invocado no combate à desinformação eleitoral, mas que, levado ao rigor técnico, seria de difícil ou impossível aplicação. A raiz do problema está numa confusão conceitual mais profunda: tratar *fake news* como sinônimo de mentira, pois o Direito não persegue mentiras em abstrato, o que o autoriza a intervir é o dano, efetivo ou potencial, a bens juridicamente protegidos, restando à mentira sem potencial lesivo um dilema no campo da ética ou moral, mas não jurídico. Disso decorre a proposta de método de verificação judicial que se organiza em dois passos sequenciais. Primeiro, verificar se o conteúdo tem potencial lesivo; sem ele, não há problema jurídico. Segundo, verificar se o conteúdo é enganoso, critério funcional que o próprio ordenamento já aplica no Direito do Consumidor e em outros institutos. Enganoso não significa necessariamente falso, pois um conteúdo pode ser composto de elementos verdadeiros e, ainda assim, induzir a conclusões incorretas. O artigo também demonstra a aplicação do método em casos concretos, discute o impacto da inteligência artificial generativa na verificação de falsidades e conclui que o Direito funciona melhor quando protege bens da vida do que quando tenta eliminar a mentira do mundo.

PALAVRAS-CHAVE: *Fake news*. Conteúdo enganoso. Potencial lesivo. Bens jurídicos tutelados. Método de verificação judicial.

ABSTRACT: This article proposes a legal understanding of fake news as deceptive content with harmful potential, taking as its starting point the opposition to the concept of “knowingly untruthful” derived from Article 323 of the Electoral Code, a concept invoked in the fight against electoral disinformation but which, when subjected to technical legal rigor, would be difficult or impossible to apply. The root of the problem

lies in a deeper conceptual confusion: treating fake news as synonymous with lying. Law does not pursue lies in the abstract; what authorizes legal intervention is harm, effective or potential, to legally protected interests. A lie devoid of harmful potential remains a dilemma in the field of ethics or morality, but not a legal one. From this follows the proposed judicial verification method, organized into two sequential steps. The first is to verify whether the content has harmful potential; without it, there is no legal problem. The second is to verify whether the content is deceptive, a functional criterion that the legal system already applies in Consumer Law and other fields. Deceptive does not necessarily mean false, since content may be composed of individually true elements and still lead to incorrect conclusions. The article also demonstrates the application of the method to concrete cases, discusses the impact of generative artificial intelligence on the verification of falsities, and concludes that Law functions better when it protects legally protected interests than when it attempts to eliminate lying from the world.

KEYWORDS: Fake news. Deceptive content. Harmful potential. Legally protected interests. Judicial verification method.

INTRODUÇÃO

O debate sobre *fake news* no Brasil produziu uma quantidade expressiva de artigos, projetos de lei e decisões judiciais, mas poucas vezes paramos para examinar se o conceito que estava sendo utilizado era o correto para os fins que se pretendia alcançar, e isso não é uma discussão abstrata, pois é o conceito que define as premissas de atuação. Acredito que tenhamos partido de uma premissa intuitiva, porém incompatível como objeto do Direito. Partimos da semântica do fenômeno acreditando que *fake news* são notícias falsas e, portanto, seriam sinônimos de mentira, cabendo ao Direito corrigir, punir e prevenir.

Assim, construiu-se sobre esse conceito uma estrutura jurídica que, embora produza resultados razoáveis em parte dos casos, revela suas fragilidades exatamente nos casos mais complexos, que, não por acaso, são os mais importantes para a proteção do processo democrático e das instituições.

O ponto de partida deste artigo é um incômodo com essa premissa, não porque ela seja inteiramente errada, mas porque ela responde à pergunta errada. Ao perguntar se um conteúdo é verdadeiro ou falso, o Direito assume para si uma função epistemológica que não tem condições de exercer com segurança e que, em uma democracia, não deveria ser sua função primária, porque a pergunta juridicamente relevante não é se o conteúdo é falso, mas

se ele causa danos ou ameaça de danos a bens juridicamente protegidos, e é exatamente essa mudança de eixo que este artigo propõe.

Essa correção conceitual tem consequências diretas na estrutura dos tipos penais aplicáveis, na responsabilidade civil das plataformas digitais, na jurisprudência eleitoral consolidada ao longo dos últimos anos e, sobretudo, nos limites constitucionais entre o combate legítimo à desinformação e a proteção da liberdade de expressão, tensão que não se resolve pela escolha de um polo em detrimento do outro, mas pela precisão do conceito que se utiliza para delimitar o campo de atuação do ordenamento.

Conceitos mal formulados produzem resultados imprevisíveis, às vezes punindo comportamentos que não deveriam ser punidos, às vezes deixando sem tutela situações que claramente a exigiriam, e quase sempre gerando a insegurança jurídica que mais compromete a credibilidade das instituições nos casos que mais exigem clareza.

A pesquisa que sustenta este artigo começou em 2016 e foi publicada sistematicamente em 2018, num momento em que o debate brasileiro ainda era amplamente dominado pela ideia de que o problema se resolveria com mais uma lei ou com a aplicação mais rigorosa dos dispositivos já existentes, especialmente do artigo 323 do Código Eleitoral, cuja redação parecia, à primeira vista, encaixar-se no problema das eleições de 2018 com uma precisão que logo se revelaria ilusória.

A proposta era diferente: antes de discutir quais instrumentos o Direito deveria mobilizar contra as *fake news*, era preciso discutir o que o Direito tem condições reais de fazer, e essa pergunta prévia levou ao conceito de *fake news* como conteúdo enganoso com potencial lesivo, cuja defesa aprofundada, com suas implicações práticas e metodológicas, este artigo desenvolve.

O argumento percorre aqui cinco etapas que se organizam em progressão lógica, iniciando pela análise das razões pelas quais o artigo 323 do Código Eleitoral, com seu critério do “fato sabidamente inverídico”, não resolve o problema que foi convocado a resolver; o exame da pergunta central que reorganiza todo o debate, a saber, se o Direito pune a mentira ou o dano e qual a relação entre um e outro; a apresentação do conceito jurídico de desinformação (*fake news*) como o conteúdo enganoso com potencial lesivo e das vantagens que ele oferece sobre a dicotomia verdade/mentira; a proposição de um método de verificação judicial organizado em dois passos sequenciais, com aplicação a casos concretos; e, por fim, a discussão sobre o impacto da inteligência artificial generativa na verificação de falsidades e sobre os bens jurídicos que o conceito busca proteger.

O argumento central deste artigo, que percorre todas essas etapas, pode ser enunciado desde o início: o Direito funciona melhor quando protege bens da vida do que quando tenta eliminar a mentira do mundo, e

essa afirmação não é cética quanto às possibilidades do ordenamento, mas, sim, mais precisa quanto ao que ele foi construído para fazer e faz bem quando opera com os conceitos adequados, como se demonstrará a seguir.

1. O PONTO DE PARTIDA: O QUE O ARTIGO 323 DO CÓDIGO ELEITORAL TENTOU RESOLVER E POR QUE NÃO RESOLVEU

Quando o debate sobre *fake news* chegou com mais força ao Brasil, mais precisamente nas eleições de 2018, o Direito Eleitoral foi convocado como a resposta natural ao desafio instaurado e, como um passe de mágica, foi logo nesse momento que o artigo 323 do Código Eleitoral foi retirado da manga em uma ação de (re)descoberta que até parecia se encaixar no problema: quem divulgar, na propaganda eleitoral, fatos que sabe inverídicos em relação a partidos ou candidatos, capazes de exercer influência perante o eleitorado, comete crime.²

O Tribunal Superior Eleitoral passou a desenvolver sua jurisprudência em torno desse dispositivo, mesmo que não o aplicasse, isto é, passou a ser considerado *fake news* na seara eleitoral o fato sabidamente inverídico.

O problema é que, se formos técnicos, esse dispositivo não serve para o que se pretende que ele sirva, não por má vontade de quem o aplica, mas por razões que estão na estrutura do próprio tipo.

O primeiro problema está no elemento subjetivo, pois o tipo exige que o agente saiba que o fato é inverídico e a questão prática que surge é como demonstrar isso em juízo. Esse grau de certeza sobre o estado mental de outra pessoa é excepcionalmente difícil de produzir probatoriamente, ainda mais em um processo criminal com garantias da presunção de inocência, *in dubio pro reo*, e, ainda, considerando que o ônus probatório recai inteiramente sobre quem acusa.

O segundo problema é ainda mais profundo, pois o tipo exige que o fato inverídico seja capaz de exercer influência perante o eleitorado. No entanto, ninguém sabe por que alguém vota em alguém; nem a Ciência Política, com todos os seus instrumentos de pesquisa, consegue estabelecer causalidade precisa entre uma mensagem específica e uma decisão de voto. Muitas vezes, nem o próprio eleitor seria capaz de reconstituir com precisão as razões que o levaram a escolher um candidato. A exigência de demonstrar influência no eleitorado, se levada ao rigor técnico que o Direito Penal exige, tornaria o tipo praticamente inaplicável.

O resultado prático é que o artigo 323 tem sido aplicado como um amparo legal geral para o combate à desinformação eleitoral, com a

2 Ver art. 323: “Divulgar, na propaganda eleitoral, fatos que sabe inverídicos em relação a partidos ou candidatos, capazes de exercer influência perante o eleitorado” (Brasil, 1965).

jurisprudência do TSE preenchendo as lacunas que o texto da lei deixa abertas. Isso funciona em boa parte dos casos e ninguém ignora que o conceito de fato sabidamente inverídico está, até hoje, presente na memória de quem trabalha com o tema. Mas o fato de algo funcionar às vezes ou, ao menos, parecer que funciona não o torna conceitualmente adequado. E, quando o conceito é inadequado, as distorções aparecem nos casos mais complexos, que são exatamente os mais importantes.

Esse incômodo foi o ponto de partida da pesquisa que resultou no conceito desenvolvido a partir de 2016 e publicado sistematicamente em 2018.³ A percepção era de que o debate estava respondendo à pergunta errada. Todos queriam saber o que é *fake news*. Poucos paravam para perguntar: para que serve o Direito, e o que ele tem condições reais de fazer?

2. A PERGUNTA QUE MUDA TUDO: O DIREITO PUNE A MENTIRA OU O DANO?

Quando se examina o ordenamento jurídico brasileiro em busca dos casos em que ele pune a mentira, o primeiro instinto é pensar na calúnia, no falso testemunho, na falsidade documental, na falsidade ideológica etc., pois há realmente uma série de tipos penais e ilícitos civis que parecem girar em torno da mentira.

Mas se se vai um pouco mais fundo em cada um deles, a pergunta que emerge é esta: o Direito está punindo esses comportamentos porque houve mentira ou porque houve ameaça de dano a um bem juridicamente protegido?

Pense no exemplo mais simples possível. Imagine que alguém, num dia ruim, é abordado por um conhecido que pergunta como você está e você responde que está ótimo, quando na verdade está mal. Você mentiu. Todavia, é evidente que não causou algum dano nem ameaçou algum bem jurídico, o que torna lógica a não punição ou o não tratamento dessa mentira no âmbito do Direito. Mas perceba que, se a mentira fosse automaticamente objeto de atuação jurídica, essa situação deveria acionar o ordenamento, o que evidentemente não acontece. Há um problema de lógica que o conceito tradicional não resolve: se mentira é proibida, por que algumas não o são?

A conclusão a que se chega, depois de percorrer os institutos do ordenamento, é que a mentira que interessa ao Direito é aquela acompanhada de potencial lesivo. A mentira sem potencial lesivo pode ser um problema ético, moral, filosófico, religioso, cultural, familiar, literário ou de qualquer outra ordem, mas não é um problema jurídico. O que o Direito persegue não é a mentira, é o dano.

3 A sistematização do conceito foi publicada pela primeira vez em: Rais, 2018.

Isso leva a uma conclusão que talvez seja a mais importante deste artigo: no sistema jurídico brasileiro, não existe um dever jurídico universal de dizer a verdade. Existe o dever de não causar dano, e o dilema entre verdade e mentira está muito mais no campo da moral e da ética do que no campo do Direito. O que está no campo do Direito é a proteção dos bens da vida, e aqui a mentira é tratada como insumo ou meio, e não como fim.

Os crimes contra a honra do Código Penal confirmam essa lógica.⁴ A calúnia é punida não porque alguém mentiu, mas porque imputar falsamente a outra pessoa fato definido como crime lesa a honra objetiva dessa pessoa, ou seja, o conceito que os outros fazem dela. O bem jurídico protegido é a honra, não a verdade. O falso testemunho é punido não porque o depoente mentiu, mas porque a mentira perante o juízo compromete a administração da justiça, que é o bem jurídico tutelado ali. Dissecado cada instituto, o que aparece sempre é o bem da vida que está sendo protegido, não a verdade em abstrato.

A mesma lógica aparece no Direito do Consumidor. O artigo 37, parágrafo 1º, do Código de Defesa do Consumidor⁵ define publicidade enganosa como aquela capaz de induzir em erro o consumidor, independentemente de ser inteira ou parcialmente falsa ou de agir por omissão. O critério não é a falsidade da afirmação, é a capacidade de induzir a erro, por isso o ordenamento já opera com essa lógica funcional há décadas, e sem grande controvérsia, mas caímos na cortina de fumaça de igualar *fake news* à mentira, e não o é, ao menos para fins de interpretação jurídica.

Feita essa travessia, o conceito de *fake news* precisava ser reposicionado. Não como notícia falsa, que é a tradução literal que cria mais problemas do que resolve. Por isso, tenho proposto que o conceito jurídico de *fake news* deveria ser: o conteúdo enganoso com potencial lesivo. Talvez, aliás, a melhor tradução não seja notícia falsa, mas notícia fraudulenta, pois a fraude, e não a falsidade, é o objeto mais próximo do Direito nesse campo, como tenho defendido desde 2018 (Canário, 2018).

3. FAKE NEWS COMO CONTEÚDO ENGANOSO COM POTENCIAL LESIVO

O conceito de *fake news* como conteúdo enganoso com potencial lesivo foi sistematizado a partir de 2016 e publicado em 2018, passando por alguns ajustes ao longo do tempo. Na formulação original, eram três os elementos: dolo, dano e falsidade. Com o refinamento da pesquisa, o

4 Ver art. 138 a 140 (Brasil, 1940).

5 Ver art. 37, § 1º: “É enganosa qualquer modalidade de informação ou comunicação de caráter publicitário, inteira ou parcialmente falsa, ou, por qualquer outro modo, mesmo por omissão, capaz de induzir em erro o consumidor” (Brasil, 1990).

conceito foi simplificado para dois passos: a verificação da enganiosidade do conteúdo e a verificação do seu potencial lesivo. Essa simplificação não significa que o dolo desapareceu da análise, mas que ele está incorporado à própria noção de enganiosidade intencional, que pressupõe que o conteúdo foi produzido ou veiculado com o objetivo de induzir a erro.

Essa formulação tem uma vantagem prática importante sobre o binômio verdade/mentira: a enganiosidade é algo que o Direito já sabe aferir. A publicidade enganosa, os vícios redibitórios, a simulação nos negócios jurídicos, o dolo como vício do consentimento: todos esses institutos lidam com a capacidade de induzir a erro, e o ordenamento brasileiro tem ferramentas desenvolvidas para fazer esse juízo. A veracidade, por outro lado, implica um juízo epistêmico que o Direito não tem condições de fazer com segurança e que, em democracia, não deveria ser função sua.

A distinção entre enganoso e falso não é meramente terminológica. Um conteúdo pode ser composto de elementos individualmente verdadeiros e, ao mesmo tempo, ser profundamente enganoso, se apresentado fora de contexto, editado seletivamente ou estruturado de forma a conduzir o receptor a uma conclusão incorreta.

Durante a pandemia de COVID-19, circularam vídeos mostrando pessoas desmaiando logo após receber vacinas, com legendas sugerindo que as imunizações causavam mortes súbitas. Os vídeos eram autênticos e as quedas aconteceram de verdade, e, realmente, as pessoas tinham acabado de ser vacinadas. Analisando cada elemento isolado, era verdadeiro, mas o conjunto era profundamente enganoso: os desmaios resultavam de reações vasovagais, fenômeno comum causado por ansiedade e pelo estresse do procedimento, sem qualquer relação com efeitos adversos das vacinas. A apresentação descontextualizada levou uma quantidade enorme de pessoas a concluir erroneamente que as vacinas eram perigosas.

Se o critério jurídico fosse a falsidade, esse conteúdo seria difícil de enquadrar, porque nenhum elemento específico era falso. Com o critério da enganiosidade, a questão se torna mais clara: o conteúdo, no contexto em que foi apresentado, induzia a uma conclusão incorreta sobre a segurança das vacinas, e essa conclusão incorreta tinha potencial de causar danos à saúde pública, ao desestimular a vacinação.

Também é possível a situação inversa: uma mentira que não engana. O exagero retórico, a hipérbole, a sátira, a paródia: todos utilizam deliberadamente afirmações que o receptor sabe não serem literalmente verdadeiras. Esses conteúdos mentem em sentido estrito, mas não enganam, porque o código comunicativo compartilhado entre emissor e receptor deixa claro que não se trata de relato factual. Para o Direito, o que importa é a enganiosidade, não a falsidade.

O segundo elemento do conceito é o potencial lesivo, e ele é o mais decisivo dos dois, pois conteúdo enganoso sem potencial lesivo não é um problema jurídico, embora possa ser um problema de outra ordem, como um problema moral ou ético, mas, por si só, não tem a capacidade de acionar a tutela do ordenamento jurídico.

Afirmar isso não é dizer que o Direito é indiferente à desinformação, é dizer que o Direito é indiferente à desinformação inofensiva e que sua intervenção se justifica quando há ameaça ou lesão a bens da vida juridicamente protegidos. Isso é consistente com a função que o Direito exerce em todas as outras áreas: ele não é um instrumento de moralização geral, mas de proteção de interesses juridicamente tutelados.

O potencial lesivo pode se manifestar em diferentes graus e direções, pode ser abstrato ou concreto, individual ou coletivo, difuso ou homogêneo. Pode, por exemplo, se dirigir à honra de uma pessoa, à integridade do processo eleitoral, à saúde pública, à segurança coletiva, às instituições democráticas etc. O que é indispensável é que haja, em algum grau verificável, a ameaça ou a lesão a algum desses bens, e a extensão e gravidade da lesão ou ameaça de lesão é que deve nortear a responsabilidade do agente.

Acrescente-se que o dano pode ser efetivo ou em potencial, e essa ressalva importa porque a natureza viral do conteúdo digital torna a lesão muitas vezes irreversível: uma vez propagada em escala, a desinformação não se apaga com a mesma velocidade com que se espalhou e, muitas vezes, aguardar o dano consumado antes de agir pode tornar a tutela jurídica praticamente inútil. Por isso, a ameaça de lesão já é suficiente para acionar o conceito.

4. OS DOIS PASSOS DO MÉTODO DE VERIFICAÇÃO JUDICIAL E ANÁLISE DE CASOS CONCRETOS

O conceito de conteúdo enganoso com potencial lesivo não é apenas uma definição, é um método que tem uma sequência lógica que orienta a verificação judicial diante dos casos concretos.

A primeira pergunta que se deve fazer diante de qualquer conteúdo supostamente desinformativo é: esse conteúdo tem potencial lesivo? Se a resposta for negativa, o processo de análise se encerra aí, pois não haverá problema jurídico a ser resolvido. O conteúdo pode ser falso, pode ser enganoso, pode ser eticamente reprovável, mas, sem potencial lesivo, não é objeto de tutela jurídica.

Isso pode parecer óbvio quando dito assim, mas, na prática cotidiana do Direito Eleitoral e da responsabilidade civil digital, há muitos casos em que essa pergunta não está sendo feita. Conteúdos com grau mínimo ou nenhum potencial de causar dano real chegam ao Judiciário e são tratados com

a mesma seriedade que casos de campanhas organizadas de desinformação com impacto massivo e muitas vezes industrial, e esse primeiro passo do método serve exatamente para separar esses casos.

Confirmado o potencial lesivo, a segunda pergunta é: o conteúdo é enganoso? Enganoso, como já argumentado anteriormente, não significa literalmente algo falso. Significa que o conteúdo, no contexto em que foi produzido e veiculado, tende a criar no receptor uma impressão distorcida da realidade.

Essa verificação leva em conta o contexto integral: o texto, a imagem, a edição, a legenda, o veículo, o público-alvo, o momento em que o conteúdo circula etc. É evidente que há um grau de subjetividade nessa avaliação, mas o mesmo grau de subjetividade está presente em praticamente toda a atividade jurídica.

O Direito Penal avalia o dolo, que é um estado mental, enquanto o Direito Civil avalia a boa-fé, que é um standard de comportamento; o Direito do Consumidor avalia a capacidade de induzir em erro o consumidor, que é um julgamento sobre percepção. A enganiosidade não é mais subjetiva do que esses outros critérios; pode ser, na prática, menos subjetiva do que tentar demonstrar se alguém sabia que determinada informação era falsa.

Há uma analogia que ajuda a entender o problema conceitual que este artigo busca resolver: um martelo é uma boa ferramenta para pregar pregos, mas não é uma boa ferramenta para desaparafusar, e isso não faz do martelo um instrumento bom ou ruim em sentido absoluto, faz dele adequado ou inadequado para fins específicos. Se alguém tenta usar um martelo para tirar um parafuso minúsculo de um controle remoto, o resultado provável é a destruição do controle, e, perceba, o problema não é o martelo, é o uso inadequado dele.

O Direito tem esse mesmo problema quando é convocado a eliminar a mentira do mundo. Não é que o Direito seja um instrumento ruim, mas é que eliminar a mentira do mundo, se isso for possível, com certeza não é uma função para a qual o ordenamento jurídico foi concebido, pois o Direito não tem os instrumentos adequados para cumprir essa tarefa. Quando se tenta usá-lo assim, o efeito é parecido com o do martelo no parafuso: a ferramenta não cumpre a função pretendida, e o objeto acaba destruído, no caso, a liberdade de expressão, a segurança jurídica, a confiança nas instituições.

Se o Direito focar sua atuação no que efetivamente pode fazer, que é proteger bens da vida contra lesões concretas ou potenciais, ele funcionará bem e será percebido como útil. Se for convocado a ser o árbitro da verdade, fracassará nessa tarefa e, o que é pior, poderá causar danos colaterais graves ao tentar cumpri-la.

Enfim, a escolha do conceito jurídico correto de *fake news* é, portanto, também uma escolha sobre os limites e as capacidades do Direito como instrumento de regulação social.

Um dos problemas que o método de verificação ajuda a resolver é o enquadramento jurídico de perguntas, não de afirmações. Há casos em que o Judiciário tem sido chamado a avaliar se um conteúdo em forma de pergunta constitui desinformação passível de sanção. Esse enquadramento é logicamente impossível, e o conceito desenvolvido aqui ajuda a explicar por quê.

Uma pergunta não afirma. E, se não afirma, não pode ser verdadeira falsa em sentido estrito, e isso não é opinião doutrinária, é uma questão de lógica elementar: o valor de verdade só se aplica a proposições, e perguntas em si não são proposições. Tentar classificar uma pergunta como fato inverídico, nos termos do artigo 323 do Código Eleitoral, é uma impossibilidade lógica que o texto do dispositivo evidentemente não contemplou.

O conceito de conteúdo enganoso com potencial lesivo lida melhor com esses casos. A pergunta, dependendo do contexto, pode ser enganosa: se estruturada de forma a pressupor como verdadeiro algo que não o é, induzindo o receptor a assumir uma premissa falsa, pode configurar conteúdo enganoso. Um exemplo claro seria uma pergunta do tipo: “Será que o candidato roubou?”, enquanto mostra uma imagem real do candidato em suas férias em um lugar paradisíaco. Perceba a enganosidade do conteúdo criando a existência de um comportamento que talvez não exista, pois o problema aqui não é a forma interrogativa, é a premissa enganosa embutida, e é claro que esse julgamento depende de análise contextual, não de enquadramento automático em categoria predefinida.

Nas eleições municipais de São Paulo de 2024, um episódio concreto ilustra bem a aplicação do método e, ao mesmo tempo, os problemas que surgem quando o conceito utilizado é inadequado.

Durante a campanha, o então candidato Ricardo Nunes chamou sua adversária Tabata Amaral de Barbie e, em resposta, a campanha de Tabata Amaral produziu um vídeo em que a candidata, numa rua, perguntava a transeuntes quem era Ricardo Nunes e ouvia como resposta: quem? A resposta de desconhecimento era então acompanhada de uma animação mostrando uma cabeça estilizada de Nunes sobre o corpo do boneco Ken, fazendo uma dancinha.

O vídeo chegou à Justiça Eleitoral e o conteúdo foi avaliado e, aplicando os critérios do conceito proposto, a conclusão é que ele não constitui desinformação passível de sanção. A manipulação é evidente e declarada: ninguém que assiste ao vídeo acredita que está vendo Ricardo Nunes dançar de verdade. O conteúdo não engana ninguém e dizer que

uma pessoa é desconhecida, ainda que pelo recurso ao humor, tampouco configura ofensa à honra objetiva que justifique intervenção jurídica.

5. O IMPACTO DA INTELIGÊNCIA ARTIFICIAL GENERATIVA E OS BENS JURÍDICOS ENVOLVIDOS

O avanço da inteligência artificial generativa adicionou uma dimensão nova ao problema da desinformação que o método de verificação judicial precisa incorporar. Há conteúdos produzidos com nível de sofisticação tal que, muitas vezes, não se consegue determinar com segurança se são ou não sintéticos, e esse cenário tem consequências diretas para o método. Se o critério fosse a falsidade, casos como esse seriam intratáveis judicialmente: sem prova pericial conclusiva de que o conteúdo é falso, o enquadramento no conceito ficaria comprometido. Mas, com o critério da enganiosidade com potencial lesivo, o problema se apresenta de forma diferente. A pergunta não é apenas se o conteúdo é tecnicamente fabricado, mas se ele, no contexto em que circulou, criou impressão incorreta da realidade com potencial de causar dano.

Isso não elimina os desafios probatórios, mas os reorienta, fazendo a discussão passar de uma questão técnica binária (é falso ou não é) para uma análise contextual mais ampla, que inclui a percepção que o conteúdo gerou, o comportamento que induziu e os bens jurídicos que ameaçou. É uma avaliação mais complexa, mas também mais honesta sobre o que o Direito tem condições de fazer.

O potencial lesivo do conteúdo enganoso pode se dirigir a uma grande variedade de bens jurídicos, e identificá-los com precisão é parte do método de verificação. O bem jurídico ameaçado determina a gravidade da lesão, a proporcionalidade da resposta e o ramo do Direito mais adequado para lidar com o caso.

A honra e a reputação de pessoas identificadas são os bens jurídicos mais diretamente atingidos por *fake news* dirigidas a indivíduos específicos. Quando um conteúdo enganoso imputa falsamente a alguém conduta criminosa, imoral ou vexatória, lesa a honra objetiva da vítima e pode produzir dano material e moral mensuráveis. Os crimes contra a honra do Código Penal e a responsabilidade civil já fornecem instrumentos adequados para essa situação.

A integridade do processo eleitoral é bem jurídico cuja proteção o ordenamento trata como prioritária em razão da proteção da integridade eleitoral, autonomia da vontade do eleitorado e isonomia entre candidaturas, afinal, o eleitorado que forma sua convicção política com base em conteúdo enganoso não exerce plenamente sua autonomia

política, e o resultado eleitoral pode não refletir a vontade real do eleitorado.

A saúde pública é outro exemplo de bem jurídico cuja relevância ficou dramaticamente evidente durante a pandemia de COVID-19. Desinformação sobre vacinas, medicamentos e condutas preventivas pode ter efeitos letais em escala: redução das taxas de vacinação, adoção de tratamentos ineficazes ou prejudiciais, resistência a medidas sanitárias necessárias. A lesão não atinge uma pessoa, mas potencialmente toda a população.

A segurança pública também pode ser comprometida por desinformação que incita violência, identifica falsamente pessoas como autoras de crimes ou distorce informações sobre eventos em andamento. Casos de violência motivados por boatos circulando em aplicativos de mensagens mostram que esse bem jurídico pode ser diretamente atingido com consequências graves e imediatas.

A identificação precisa do bem jurídico ameaçado serve também para determinar a gradação do potencial lesivo e, com isso, orientar a proporcionalidade da resposta jurídica. Um conteúdo que ameaça a saúde pública em escala nacional não deve receber o mesmo tratamento que um conteúdo que lesiona a honra de um indivíduo em contexto de alcance limitado.

Uma das vantagens do conceito proposto é que ele não exige a criação de institutos jurídicos inteiramente novos, ele se integra de forma natural ao sistema já existente, aproveitando os instrumentos que o ordenamento brasileiro já desenvolveu para situações análogas.

No Direito Penal, o conceito de conteúdo enganoso com potencial lesivo é compatível com os crimes contra a honra já existentes. A calúnia, que exige a imputação falsa de fato definido como crime, incorpora tanto a enganosidade da imputação quanto o potencial lesivo à honra objetiva. O falso testemunho protege a administração da justiça. O estelionato protege o patrimônio contra a fraude. Em todos esses casos, a estrutura do tipo já contempla, implicitamente, os dois elementos do conceito proposto.

No Direito Eleitoral, o artigo 323 do Código Eleitoral e as resoluções do TSE, especialmente a Resolução n. 23.610/2019, aproximam-se do conceito ao exigir, além da falsidade, a capacidade de influenciar o eleitorado. O problema do dispositivo, como apontado no início deste capítulo, está na dificuldade probatória de demonstrar ambos os elementos com o rigor que o Direito Penal exige. O conceito proposto oferece um caminho alternativo que pode ser mais operacionalizável na prática.

Na responsabilidade civil, o sistema brasileiro já dispõe de todos os instrumentos necessários para tratar a desinformação. A tríade conduta ilícita, dano e nexo causal é suficiente para enquadrar os casos de *fake news*

com potencial lesivo. A novidade, no campo do dano, é o reconhecimento do dano social⁶ como categoria aplicável quando a desinformação em larga escala rebaixa o nível de confiança coletiva ou compromete bens de natureza difusa, como a saúde pública ou a integridade do processo democrático.

O julgamento do Tema 987 pelo Supremo Tribunal Federal (STF) em 2025⁷ aproximou-se dessa lógica ao flexibilizar o artigo 19 do Marco Civil da Internet para permitir a responsabilização de plataformas por conteúdos gravemente lesivos a bens jurídicos fundamentais. Por mais que haja muitas críticas a respeito da decisão, o STF decidiu que o modelo de imunidade ampla das plataformas não poderia funcionar como escudo para condutas que produzem dano grave a bens de alta relevância, o que é consistente com a premissa de que o potencial lesivo é o elemento que justifica a intervenção do Direito.

No plano constitucional, a tese de que o combate à desinformação lesiva é incompatível com a liberdade de expressão inverte a lógica que o STF tem aplicado de forma consistente. No HC 82.424,⁸ o Tribunal reconheceu que discurso de ódio não está protegido constitucionalmente, enquanto na ADPF 130⁹ vedou a censura prévia, mas reconheceu a legitimidade da responsabilização posterior, e, na ADI 7261, validou remoção célere de desinformação eleitoral. O que o STF tem dito, em diferentes contextos, é que a liberdade de expressão não protege manifestações que causam lesão a bens jurídicos fundamentais. Isso não é inovação, é a aplicação de princípio que sempre esteve presente no ordenamento.

Mas talvez aqui outro ganho possa ser trazido pelo conceito jurídico de desinformação como conteúdo enganoso com potencial lesivo. Diante do debate sobre a liberdade de expressão e a desinformação, muito se discute sobre os limites da verdade e da mentira, mas jamais discutimos um “direito de enganar”, provocando ou ameaçando um bem jurídico de lesão, como se isso fosse derivado da liberdade de expressão. Aqui, mais uma vez, parece que o conceito jurídico de desinformação apresenta vantagens ao sabidamente inverídico.

6 O Enunciado 456 da V Jornada de Direito Civil do Conselho da Justiça Federal (2012) consolidou o dano social como espécie implícita no art. 944 do Código Civil (Azevedo, 2004).

7 O STF flexibilizou o art. 19 do Marco Civil da Internet para permitir a responsabilização de plataformas por conteúdos gravemente lesivos a bens jurídicos fundamentais (Brasil, 2025).

8 O STF inaugurou esse entendimento no HC 82.424/RS (Caso Ellwanger), julgado em 17 de setembro de 2003, ao reconhecer que o discurso de ódio não está protegido constitucionalmente porque atenta contra a dignidade da pessoa humana e o princípio da igualdade.

9 O STF vedou a censura prévia, mas reconheceu a legitimidade da responsabilização posterior por manifestações que violem direitos fundamentais (Brasil, 2009).

CONCLUSÃO

O ponto de partida deste artigo foi um incômodo conceitual com o “sabidamente inverídico” como conceito de desinformação, sobretudo derivado do artigo 323 do Código Eleitoral, e com o modo como o debate brasileiro sobre desinformação ficou preso à dicotomia verdade/mentira. O percurso argumentativo chegou a uma proposta que, acredita-se, possa resolver o problema com mais consistência: *fake news* devem ser compreendidas juridicamente como conteúdo enganoso com potencial lesivo.

Essa proposta não é apenas uma simplificação, mas, sim, uma proposta de correção de rota. O Direito não tem condições de eliminar a mentira do mundo, e tentar colocá-lo nessa posição é usar a ferramenta errada para o trabalho errado. O que o Direito tem condições de fazer, e faz bem quando os conceitos são adequados, é proteger bens da vida contra lesões concretas ou potenciais, e é esse o espaço que considero correto para o debate sobre desinformação online.

O método de verificação que o conceito gera é operacionalizável na prática judicial: verificar primeiro o potencial lesivo, verificar depois a enganiosidade. Se o conteúdo não passa pelo primeiro filtro, não é um problema jurídico. Se passa pelos dois, configura ilícito passível de tutela pelo ordenamento, nos diferentes ramos que já dispõem de instrumentos para tanto.

Os casos concretos mostram que o método funciona nos dois sentidos: ele enquadra adequadamente as situações que merecem tutela jurídica e, ao mesmo tempo, deixa de fora as que não merecem, como a pergunta que não afirma, a sátira que não engana, o humor que não lesiona. Isso é tão importante quanto o primeiro resultado: um conceito que inclui o que deve incluir e exclui o que deve excluir é um conceito juridicamente útil.

O avanço da inteligência artificial generativa torna essa discussão ainda mais urgente, sendo necessário enfrentar a alternativa entre usar o Direito como árbitro da verdade ou como protetor dos bens jurídicos.

REFERÊNCIAS

AZEVEDO, Antônio Junqueira de. Por uma nova categoria de dano na responsabilidade civil: o dano social. *Revista Trimestral de Direito Civil*, Rio de Janeiro, ano 5, n. 19, p. 211-218, jul./set. 2004.

BRASIL. Código Eleitoral. Lei nº 4.737, de 15 de julho de 1965. *Diário Oficial da União*, Brasília, DF, 1965.

BRASIL. Código Penal. Decreto-Lei nº 2.848, de 7 de dezembro de 1940. *Diário Oficial da União*, Brasília, DF, 1940.

BRASIL. Código de Defesa do Consumidor. Lei nº 8.078, de 11 de setembro de 1990. *Diário Oficial da União*, Brasília, DF, 1990.

BRASIL. Conselho da Justiça Federal. *V Jornada de Direito Civil*. Enunciado 456. Brasília: CJP, 2012.

BRASIL. Supremo Tribunal Federal. *RE 1.037.396 (Tema 987)*. Relator: Min. Dias Toffoli. Julgamento: 26 jun. 2025. Brasília, DF: STF, 2025.

BRASIL. Supremo Tribunal Federal. *ADPF 130/DF*. Relator: Min. Carlos Ayres Britto. Julgamento: 30 abr. 2009. Brasília, DF: STF, 2009.

CANÁRIO, Pedro. “A melhor tradução para *fake news* não é notícia falsa, é notícia fraudulenta”. *Consultor Jurídico*, 2018. Disponível em: <https://www.conjur.com.br/2018-ago-12/entrevista-diogo-raiz-professor-direito-eleitoral/>. Acesso em: 12 fev. 2026.

RAIS, Diogo (coord.). *Fake News: a conexão entre a desinformação e o Direito*. 1. ed. São Paulo: Thomson Reuters Brasil, 2018.