

REGULAÇÃO DA DESINFORMAÇÃO NO AMBIENTE DIGITAL: DESAFIOS DEMOCRÁTICOS, PROTEÇÃO DE DADOS E INTELIGÊNCIA ARTIFICIAL

REGULATING DISINFORMATION IN THE DIGITAL ENVIRONMENT: DEMOCRATIC CHALLENGES, DATA PROTECTION, AND ARTIFICIAL INTELLIGENCE

Luciano Henrique Caixeta Viana ¹

Simone Tavares de Andrade²

RESUMO

O presente artigo examina os principais desafios associados à regulação da desinformação no ambiente digital, com ênfase nos efeitos das fake news sobre o regime democrático, a formação da opinião pública e a credibilidade das instituições. A investigação propõe-se a discutir um modelo regulatório em rede, com foco na responsabilização das plataformas digitais, na promoção da transparência nos processos de moderação de conteúdo e na salvaguarda da liberdade de expressão. Para tanto, adota-se abordagem metodológica de natureza qualitativa, com base em pesquisa bibliográfica e análise documental de marcos normativos nacionais e internacionais, a exemplo da Lei Geral de Proteção de Dados (LGPD) e do Digital Services Act (DSA), à luz das diretrizes da Agenda 2030. Os resultados apontam que a adoção de um modelo descentralizado de regulação informacional revela-se uma alternativa promissora, desde que estruturada com mecanismos de fiscalização robustos e orientada por princípios que equilibrem inovação tecnológica e proteção de direitos fundamentais. Conclui-se que a governança digital contemporânea deve incorporar a atuação colaborativa de múltiplos agentes, estatais, privados e da sociedade civil, e assegurar tanto a proteção de dados pessoais quanto o uso ético e transparente da Inteligência Artificial nos processos de moderação de conteúdos.

Palavras-chave: Desinformação digital. Regulação em rede. Liberdade de expressão. Proteção de dados. Inteligência Artificial.

¹ Doutorando em Administração Pública pela FEA-RP/USP, mestre e bacharel em Direito pela FDRP/USP. Pesquisador em Direito Constitucional, Tributário e Direitos Humanos. Integra o Centro de Pesquisa em Gestão e Políticas Públicas Contemporâneas (GPublic/IEA-USP). E-mail: caixeta.viana@gmail.com

² Mestre em Direito (FDRP - USP). Pós-Graduada em Direito e Processo do Trabalho (UNIDERP). Graduada em Direito (UNIUBE). Graduada em Licenciatura em Letras Português e Inglês (UnicV). Graduada em Tecnólogo em Tutoria EAD (FCE Faculdade Campos Elíseos). Líder do Grupo de Pesquisa A Transformação do Direito do Trabalho na Sociedade Pós-Moderna e Seus Reflexos no Mundo do Trabalho (GEDTRAB) - FDRP USP. Integrante do Núcleo de estudos "O Trabalho Além do Direito do Trabalho" - TADT FD-USP. Participante do Laboratório de Letramento da FFLCH/USP. E-mail: academico.simone@gmail.com Currículo Lattes: <http://lattes.cnpq.br/4108252466134860> ORCID ID: <https://orcid.org/0000-0003-1958-0538>.

ABSTRACT

This article examines the main challenges associated with regulating disinformation in the digital environment, with an emphasis on the effects of fake news on the democratic regime, the formation of public opinion, and the credibility of institutions. The research proposes to discuss a network regulatory model, focusing on the accountability of digital platforms, the promotion of transparency in content moderation processes, and the safeguarding of freedom of expression. To this end, a qualitative methodological approach is adopted, based on bibliographic research and documentary analysis of national and international regulatory frameworks, such as the General Data Protection Law (LGPD) and the Digital Services Act (DSA), in light of the guidelines of Agenda 2030. The results indicate that the adoption of a decentralized model of information regulation is a promising alternative, provided that it is structured with robust oversight mechanisms and guided by principles that balance technological innovation and the protection of fundamental rights. It is concluded that contemporary digital governance must incorporate the collaborative action of multiple agents, whether state, private, or civil society, and ensure both the protection of personal data and the ethical and transparent use of Artificial Intelligence in content moderation processes.

Keywords: Digital disinformation. Network regulation. Freedom of expression. Data protection. Artificial Intelligence.

1 INTRODUÇÃO

A disseminação de notícias falsas configura um dos mais complexos contextos regulatórios e democráticos contemporâneos no contexto da era digital. Tendo como ponto de partida esse cenário, observa-se que o desenvolvimento acelerado das tecnologias de informação e comunicação, aliado à ampliação do alcance das redes sociais, estabeleceu um ambiente propício à difusão massiva de conteúdos inverídicos e manipulados, com impactos diretos sobre a formação da opinião pública e a estabilidade das instituições democráticas.

A partir dessa constatação, a propagação sistemática de desinformação compromete a qualidade do debate público, influencia negativamente a tomada de decisão dos cidadãos, interfere em processos eleitorais e fragiliza a credibilidade dos meios de comunicação tradicionais. Trata-se, portanto, de um fenômeno que transcende a esfera individual, demandando respostas institucionais articuladas e eficazes diante de sua natureza estrutural e sistêmica.

Neste sentido, a lógica de funcionamento das plataformas digitais favorece a ampla circulação de conteúdos enganosos, impulsionada por algoritmos que privilegiam o engajamento e a viralização em detrimento da acurácia informacional. Dessa forma, esse mecanismo intensifica a polarização discursiva e a disseminação de narrativas artificiais, dificultando a distinção entre fatos verificáveis e construções deliberadamente falseadas com intuito de manipulação da opinião pública.

À luz do exposto, a regulação da desinformação digital desponta como tema central nos debates contemporâneos sobre governança da internet, liberdade de expressão e segurança informacional. Ademais, as estratégias normativas adotadas por diferentes países oscilam entre modelos mais restritivos, como a remoção compulsória de conteúdos, e abordagens que privilegiam a transparência algorítmica e a autorregulação por parte das plataformas. No caso brasileiro, o Projeto de Lei n.º 2.630/2020, denominado “PL das *Fake News*”, propõe mecanismos de responsabilização das plataformas digitais e diretrizes para verificação de conteúdos, suscitando intensas discussões quanto a seus efeitos sobre a liberdade de expressão e a regulação do ecossistema informacional.

No que tange ao panorama internacional, a União Europeia instituiu o *Digital Services Act* (DSA), impondo às grandes empresas de tecnologia regras mais rígidas relativas à moderação de conteúdos e à transparência nos processos decisórios internos. Por sua vez, nos Estados Unidos, as discussões concentram-se na eventual reformulação da Seção 230 do *Communications Decency Act*, que atualmente concede imunidade às plataformas digitais em relação ao conteúdo gerado por terceiros, ensejando debates sobre os limites da responsabilidade das corporações tecnológicas.

Diante desse contexto complexo, destaca-se a proposta apresentada pela Agência Nacional de Telecomunicações (Anatel), que visa à implementação de um modelo de regulação em rede, alicerçado na atuação cooperativa entre agências reguladoras, empresas tecnológicas e entidades da sociedade civil. Tal proposta, por conseguinte, objetiva alcançar um equilíbrio entre o combate à desinformação e a preservação das liberdades comunicativas, por meio de um arranjo regulatório descentralizado, dinâmico e responsivo.

Isto posto, este artigo propõe uma análise crítica das principais situações vinculadas à adoção desse modelo regulatório, explorando suas implicações jurídicas, tecnológicas e políticas. Além disso, investiga-se a interseção entre a responsabilidade das plataformas digitais, os mecanismos de transparência na moderação de conteúdo e a necessidade de salvaguardar um espaço público informacional alicerçado em dados verificáveis e confiáveis.

2 DESENVOLVIMENTO

2.1 OBSTÁCULOS E PERSPECTIVAS DE UM MODELO DE REGULAÇÃO EM REDE NO ENFRENTAMENTO DAS FAKE NEWS

O fenômeno da desinformação, materializado na disseminação de notícias falsas, configura uma das expressões mais recorrentes e preocupantes no contexto digital contemporâneo. Caracteriza-se pela veiculação deliberada de conteúdos falsos ou enganosos, com o propósito de manipular percepções sociais e condutas coletivas.

Nesse sentido, conforme destacam Gasparetto, Pedroso e Oliveira (2022), a proliferação dessas práticas informacionais compromete a integridade do debate público e fragiliza os pilares democráticos. Diante desse cenário, observa-se que a desinformação compromete a percepção da realidade dos cidadãos, com impactos relevantes sobre os processos eleitorais, as políticas de saúde pública e a credibilidade das instituições de comunicação.

Na esteira dessa constatação, as redes sociais se configuram como ambiente privilegiado para a propagação de conteúdos inverídicos, dado que operam com elevada velocidade e ausência de mecanismos prévios de verificação (CASTELLS, 2013). Tal constatação evidencia que essa dinâmica favorece a amplificação de narrativas distorcidas, gerando externalidades negativas que vão desde a desarticulação do espaço público deliberativo até a intensificação de discursos discriminatórios e violentos.

Para que seja possível avançar nas relevantes discussões acerca das questões aqui elencadas, é imprescindível considerar que um dos principais efeitos colaterais da disseminação de fake news reside na difusão do discurso de ódio, que se manifesta por meio de expressões preconceituosas e incitadoras de violência contra grupos sociais específicos. Ainda, de acordo com Silva et al. (2011), tais discursos são compostos por elementos discriminatórios que atingem diretamente a dignidade humana e os direitos fundamentais de indivíduos e coletividades.

Sob o enfoque jurídico, é possível reconhecer que a Lei n.º 7.716/1989 criminaliza a incitação à discriminação com base em raça, religião e nacionalidade, tendo seu escopo interpretativo ampliado pelo Supremo Tribunal Federal para abranger manifestações de homofobia e transfobia, equiparando-as ao crime de racismo (BRASIL, 1989). A partir dessas considerações, chega-se inevitavelmente à constatação de que a delimitação entre liberdade de expressão e discurso de ódio constitui um obstáculo regulatório persistente, exigindo uma abordagem normativa sensível e tecnicamente refinada.

Nesse sentido, vale ressaltar, ainda, que as grandes plataformas digitais exercem um papel ambivalente no ecossistema informacional: ao mesmo tempo em que atuam como vetores da disseminação de conteúdo, também são responsáveis por adotar mecanismos de moderação voltados à contenção de abusos e desinformações (GASPARETTO; PEDROSO;

OLIVEIRA, 2022). Em outras palavras, a atual arquitetura algorítmica prioriza conteúdos sensacionalistas e polarizadores, uma vez que tais materiais promovem maior engajamento e resultam em ganhos publicitários mais expressivos.

Outrossim, para além do exposto, faz-se necessário lançar luz sobre o fato de que a formulação de políticas eficazes de moderação de conteúdo inevitavelmente suscita tensões com os princípios da liberdade de expressão, especialmente diante de possíveis alegações de censura e de excessiva delegação da regulação do discurso público a entes privados (DONEDA; ARANHA, 2020).

Ademais, outros aspectos considerados referem-se à estratégia regulatória adotada por determinadas plataformas, que inclui parcerias com agências de verificação de fatos e a exclusão de conteúdos prejudiciais. Contudo, não obstante o exposto, a ausência de critérios claros e uniformes na aplicação dessas medidas evidencia a necessidade de maior transparência e aprimoramento constante (SARLET, 2019).

A título de conclusão, cabe reiterar que a regulação do conteúdo digital tem figurado como tema central nas agendas legislativas internacionais. No contexto brasileiro, o Projeto de Lei n.º 2.630/2020, denominado “PL das *Fake News*”, propõe a responsabilização das plataformas digitais pelos conteúdos disseminados e prevê a criação de um órgão independente para fiscalização do ambiente informacional (BRASIL, 2020).

Sob a ótica comparada, observa-se que a União Europeia implementou o DSA, que impõe diretrizes mais rigorosas às *big techs* no que se refere à exclusão de conteúdos prejudiciais e à transparência dos mecanismos de moderação (EUROPEAN COMMISSION, 2022). Já no cenário norte-americano, a discussão gira em torno da Seção 230 do *Communications Decency Act*, que atualmente garante imunidade às plataformas digitais pelo conteúdo gerado por terceiros, mas que tem sido objeto de intensas propostas de revisão (BATISTA; AGUIAR, 2020).

Tendo apresentado a temática exposta, necessário se faz abordar a proposta da Agência Nacional de Telecomunicações (Anatel) de um modelo de regulação em rede para o Brasil, que surge como uma alternativa inovadora para enfrentar os desafios da moderação de conteúdo. O modelo, alicerçado na cooperação entre agências reguladoras, empresas tecnológicas e organizações da sociedade civil, propõe um sistema descentralizado de controle e fiscalização (SALINAS, 2023).

Embora essa abordagem traga consigo avanços, como maior flexibilidade e redução da carga regulatória estatal, há de se pontuar que ela também enfrenta obstáculos relevantes,

sobretudo no que se refere à coordenação entre os múltiplos atores envolvidos e às assimetrias informacionais existentes entre os entes públicos e privados.

2.2 A PROTEÇÃO DE DADOS PESSOAIS NA ERA DIGITAL: DESAFIOS E PERSPECTIVAS À LUZ DA AGENDA 2030

A temática da proteção de dados pessoais assumiu papel central nas discussões contemporâneas acerca dos direitos fundamentais, especialmente diante das transformações operadas pela era digital nas relações interpessoais, comerciais e institucionais. Tendo como ponto de partida esse cenário, o crescimento exponencial da coleta, do tratamento e do armazenamento de informações por entes públicos e privados tem gerado desafios substanciais à garantia da privacidade e da segurança informacional dos indivíduos (VASCONCELOS, 2024).

Nesse sentido, a digitalização intensiva de serviços, tanto públicos quanto privados, promoveu o uso de grandes volumes de dados e algoritmos preditivos voltados à personalização de experiências e à automatização de decisões. Tal contexto intensifica os riscos de violação da privacidade e possibilita a manipulação comportamental, sobretudo por meio de estratégias de segmentação publicitária e disseminação de desinformação (HABERMAS, 2020).

A partir dessas considerações, a exploração indevida e o vazamento de dados, além de comprometerem garantias fundamentais, podem afetar negativamente a estabilidade política e social dos Estados, o que justifica a criação de marcos regulatórios aptos a mitigar tais riscos.

No contexto brasileiro, a promulgação da Lei Geral de Proteção de Dados Pessoais (Lei n.º 13.709/2018) representou um marco normativo estruturante para a segurança digital e a governança da privacidade. Inspirada no Regulamento Geral sobre a Proteção de Dados (GDPR), da União Europeia, a LGPD disciplina o tratamento de dados pessoais no país, estabelecendo princípios e diretrizes que asseguram maior transparência e conferem ao titular controle sobre suas informações (BRASIL, 2018).

Dentre os aspectos normativos centrais da LGPD, cabe salientar que destacam-se os princípios da Finalidade e Adequação, que vedam a utilização de dados para fins diversos daqueles previamente informados; o Consentimento e a Legitimidade, que exigem autorização explícita do titular, salvo exceções legais; a Segurança e a Prevenção, que impõem a adoção de medidas técnicas e administrativas para evitar incidentes; e a Transparência e os Direitos do Titular, que garantem acesso, correção e exclusão dos dados pelo indivíduo (BRASIL, 2018).

Sob essa ótica, a implementação da LGPD não apenas visa à proteção do indivíduo frente às novas tecnologias, como também contribui para a construção de um ecossistema digital mais seguro, transparente e confiável — condições imprescindíveis ao desenvolvimento sustentável e à consolidação da governança digital.

Além disso, a relevância da proteção de dados encontra respaldo na Agenda 2030 das Nações Unidas, sobretudo em seus Objetivos de Desenvolvimento Sustentável (ODS). O ODS 16, por exemplo, propõe a promoção de sociedades pacíficas, justas e inclusivas, enquanto o ODS 9 trata da criação de infraestrutura resiliente e da promoção da inovação sustentável (ONU, 2015).

Dessa forma, a regulamentação da segurança da informação e o avanço das políticas públicas voltadas à governança digital revelam-se indispensáveis à contenção dos riscos inerentes à manipulação de dados, bem como ao fortalecimento da confiança cidadã nas instituições democráticas (ONU, 2015).

No que concerne ao ODS 9, a proteção de dados se articula com a necessidade de desenvolver infraestrutura tecnológica resiliente, especialmente diante do avanço de tecnologias emergentes como Inteligência Artificial e Internet das Coisas (IoT). Esses cenários exigem modelos regulatórios robustos e responsivos, capazes de assegurar a integridade dos dados e a privacidade dos usuários, sem comprometer o potencial inovador e o crescimento econômico sustentável. Assim, além de refletir sobre os benefícios da inovação, a governança de dados deve buscar um equilíbrio entre o compartilhamento de informações e a mitigação de vulnerabilidades cibernéticas, garantindo resiliência às infraestruturas digitais (SOUZA, 2021).

Outrossim, para além do exposto, a interface entre proteção de dados e desenvolvimento sustentável manifesta-se na formulação de políticas públicas orientadas à inclusão digital e à equidade no acesso à informação. Como argumenta Castells (2013), o acesso a dados seguros e confiáveis é elemento essencial para a constituição de sociedades democráticas e igualitárias, prevenindo que populações vulneráveis sejam desproporcionalmente afetadas por práticas abusivas de coleta e uso de dados.

Portanto, considerando que reconhecer a proteção de dados como direito fundamental fortalece não apenas o regime democrático, mas também impulsiona a inovação responsável e o desenvolvimento econômico inclusivo, em consonância com os princípios da Agenda 2030.

Cabe destacar, nesse contexto, que a importância dessa discussão é corroborada por legislações como o GDPR europeu e a própria LGPD brasileira, as quais estabelecem parâmetros normativos para o tratamento responsável de dados pessoais e criam instrumentos

de transparência e controle por parte dos titulares (BRASIL, 2018; EUROPEAN COMMISSION, 2018). Ambas evidenciam a necessidade de uma governança de dados com escopo global, na qual a proteção da privacidade seja concebida não meramente como exigência legal, mas como fundamento ético para sociedades mais equitativas e sustentáveis.

Em linhas gerais, cabe reiterar que a garantia da privacidade e da segurança informacional configura condição essencial à promoção dos direitos individuais em uma era marcada pela intensa digitalização de serviços e relações. Ao assegurar mecanismos de governança e acesso transparente às informações, a LGPD também se converte em instrumento de combate às desigualdades digitais, promovendo um ambiente mais justo e acessível.

Contudo, não obstante o exposto, a efetivação da LGPD no Brasil ainda enfrenta desafios expressivos, notadamente no que tange à criação de mecanismos fiscalizatórios eficazes e à adaptação de organizações públicas e privadas ao novo paradigma normativo. A Autoridade Nacional de Proteção de Dados (ANPD), responsável pela supervisão da lei, depara-se com limitações estruturais e orçamentárias, que dificultam a plena realização de suas atribuições institucionais (VASCONCELOS, 2024).

Ainda, a conscientização social sobre a importância da privacidade e da segurança de dados permanece incipiente, o que torna premente a formulação de políticas públicas voltadas à alfabetização digital e à educação em direitos informacionais.

Paralelamente, a proliferação de desinformação na internet desponta como um dos maiores desafios regulatórios contemporâneos. O Projeto de Lei n.º 2.630/2020, conhecido como “PL das *Fake News*” propõe a responsabilização das plataformas digitais pela verificação de conteúdos e a criação de órgão fiscalizador autônomo (BRASIL, 2020). Todavia, a implementação dessa proposta suscita impasses de ordem técnica e jurídica, dada a complexidade inerente à moderação de grandes volumes de conteúdo digital, bem como os riscos potenciais à liberdade de expressão.

Neste sentido, faz-se mister salientar que a regulamentação da desinformação deve ser conduzida com cautela, primando pela transparência e pela participação democrática, a fim de evitar mecanismos de censura indevida e fomentar a educação midiática como ferramenta de resistência cidadã frente aos riscos da desinformação.

A título de conclusão, cabe reiterar que a proteção de dados pessoais e a regulação digital constituem pilares indispensáveis à preservação dos direitos fundamentais na sociedade da informação. A LGPD e os instrumentos normativos correlatos fornecem as bases para um ambiente digital equilibrado, no qual se harmonizem liberdade, segurança e transparência, valores essenciais à construção de um futuro digital ético, democrático e sustentável.

2.3A UTILIZAÇÃO DE INTELIGÊNCIA ARTIFICIAL NA GERAÇÃO AUTOMATIZADA DE NARRATIVAS JURÍDICAS

A crescente digitalização dos procedimentos judiciais e administrativos tem, paulatinamente, impulsionado a incorporação de tecnologias disruptivas no campo jurídico, com especial ênfase na Inteligência Artificial (IA). Nesse sentido, a aplicação de algoritmos e sistemas de aprendizado de máquina na esfera jurídica tem promovido avanços na automação de tarefas repetitivas e na análise preditiva de decisões, contribuindo para a otimização da tramitação processual e para o incremento da eficiência na gestão de litígios (SALINAS, 2023).

Neste cenário, observa-se que a IA tem sido amplamente empregada em atividades como pesquisa jurisprudencial, elaboração automatizada de pareceres e petições, revisão contratual e previsão de desfechos processuais, possibilitando aos operadores do Direito o acesso a informações de forma mais célere e precisa. Ademais, iniciativas como a aplicação da tecnologia blockchain na regulação de conteúdos digitais evidenciam a convergência entre inovação tecnológica e regulação jurídica, ao conferir maior transparência e rastreabilidade à fiscalização de informações disponibilizadas em ambiente online (SALINAS, 2023).

A partir dessas considerações, infere-se que a adoção da Inteligência Artificial para fins de automação de decisões judiciais suscita relevantes questionamentos de natureza técnica, ética e constitucional. A progressiva dependência de algoritmos na tomada de decisões jurídicas pode acarretar riscos concernentes à opacidade dos critérios utilizados, à possibilidade de discriminação algorítmica e à limitação da revisão humana tempestiva.

Sobreleva refletir que, segundo parte da doutrina, os modelos preditivos de decisão tendem a reproduzir padrões normativos enviesados presentes no ordenamento jurídico, o que pode acarretar a perpetuação de desigualdades estruturais e dificultar a superação de erros sistêmicos. Nessa perspectiva, a ausência de transparência nos sistemas automatizados comumente descritos como "caixas-pretas", configura obstáculo à efetividade da tutela jurisdicional, uma vez que inviabiliza o pleno conhecimento, pelas partes, dos fundamentos que orientaram a decisão (GASPARETTO; PEDROSO; OLIVEIRA, 2022).

Para que seja possível avançar na reflexão sobre essa problemática, é imperativo reconhecer que a implementação de sistemas de IA voltados à atividade decisória demanda a adoção de mecanismos rigorosos de fiscalização e controle, de modo a garantir sua conformidade com os princípios constitucionais do contraditório e da ampla defesa. A

introdução de ferramentas de explicabilidade (“*explainability*”) e de auditoria algorítmica revela-se, portanto, imprescindível para a legitimação das decisões automatizadas e a proteção de direitos fundamentais.

Outrossim, cabe salientar que a geração de narrativas jurídicas por meio de sistemas de IA levanta discussões quanto à autoria e à responsabilidade pelo conteúdo produzido. A utilização de tecnologias automatizadas na redação de petições, pareceres e contratos enseja reflexões sobre a autenticidade e validade desses documentos no ordenamento jurídico, especialmente no tocante à definição do agente responsável por seus efeitos (KOETZ, 2024).

Neste sentido, cumpre observar que o ordenamento jurídico vigente não dispõe, de forma expressa, de normativas específicas que regulem a responsabilidade decorrente de atos praticados por sistemas de IA autônomos.

Diante disso, situações em que documentos gerados por inteligência artificial contenham informações incorretas ou induzam a erro suscitam a indagação: a quem caberia a responsabilização pelos danos eventualmente ocasionados? Tal problemática demanda a construção de marcos regulatórios específicos que possibilitem a compatibilização entre inovação tecnológica e segurança jurídica.

Tendo em vista essa complexidade, observa-se que a intensificação da disseminação de conteúdos falsos e de discursos de ódio nas plataformas digitais tem estimulado propostas de regulação que incorporam o uso de IA como instrumento de monitoramento e contenção da desinformação. No cenário brasileiro, o Projeto de Lei das Fake News contempla a adoção de tecnologias voltadas à detecção e à remoção de conteúdos nocivos (SALINAS, 2023).

Entre as propostas em debate, destaca-se o modelo de regulação em rede, defendido por entidades como a Agência Nacional de Telecomunicações (Anatel), o qual prevê a constituição de um ecossistema de fiscalização compartilhada entre entes estatais e atores privados. Nessa configuração, atribui-se às *big techs* e às organizações da sociedade civil um papel ativo na moderação de conteúdos em ambiente digital. Entretanto, a transferência de tais competências a entidades privadas suscita preocupações em torno da *accountability* e dos riscos de censura algorítmica (GASPARETTO; PEDROSO; OLIVEIRA, 2022).

Diante do exposto, é necessário reconhecer que a utilização de IA na regulação de discursos em ambientes digitais deve observar, de maneira rigorosa, os princípios do Estado Democrático de Direito, de modo a evitar que os mecanismos automatizados configurem formas de controle excessivo ou impliquem em restrições indevidas à liberdade de expressão.

A título de conclusão, constata-se que o avanço da Inteligência Artificial no campo jurídico exige a formulação de um modelo regulatório equilibrado, apto a compatibilizar o progresso tecnológico com a preservação dos princípios fundamentais que orientam o ordenamento jurídico. Elementos como transparência, auditabilidade e segurança jurídica devem constituir pilares inegociáveis na adoção de sistemas de IA voltados à automação de decisões e à regulação de conteúdos digitais, assegurando que tais tecnologias sejam utilizadas de forma ética, responsável e comprometida com os valores democráticos.

3 CONCLUSÃO

A propagação de *fake news* no ambiente digital configura-se como um dos principais desafios à salvaguarda da integridade do debate público e à proteção das instituições democráticas. A partir desse cenário, conforme discutido ao longo desta análise, a velocidade e o alcance proporcionados pelas plataformas digitais intensificam a disseminação de desinformação, afetando diretamente a formação da opinião pública, a condução dos processos eleitorais, a saúde coletiva e a credibilidade dos meios de comunicação.

Diante disso, torna-se premente a formulação de um modelo regulatório capaz de compatibilizar a liberdade de expressão com a mitigação dos danos ocasionados pela circulação de conteúdos falsos. A proposta de regulação em rede, conforme explorada, desponta como alternativa inovadora ao envolver múltiplos agentes, estatais, privados e da sociedade civil, no processo de fiscalização e moderação informacional, promovendo a descentralização do controle e a ampliação da transparência na governança digital. Contudo, essa abordagem também apresenta entraves, especialmente no que se refere à articulação eficaz entre os diferentes atores envolvidos e à necessidade de garantir que os mecanismos de moderação não incorporem práticas de censura arbitrária.

Paralelamente, a proteção de dados pessoais, estreitamente vinculada à agenda regulatória digital, assume papel central no contexto contemporâneo, em razão do volume crescente de informações sensíveis que circulam no meio virtual. A Lei Geral de Proteção de Dados (LGPD), inspirada no Regulamento Geral sobre a Proteção de Dados (GDPR) da União Europeia, estabelece parâmetros normativos fundamentais para assegurar a privacidade dos indivíduos, bem como para fomentar a transparência e o controle sobre o tratamento de suas informações.

Entretanto, a efetividade dessa legislação está condicionada à existência de estruturas fiscalizatórias robustas e à consolidação de uma cultura de conscientização social quanto à importância da privacidade no ecossistema digital.

Ademais, a incorporação crescente da Inteligência Artificial nas esferas jurídica e regulatória introduz novas camadas de complexidade ao debate. A automação de decisões e a elaboração de narrativas jurídicas por meio de sistemas de IA suscitam preocupações relativas à opacidade dos processos decisórios, à responsabilização pelos atos automatizados e à possibilidade de reprodução de vieses discriminatórios nos modelos algorítmicos.

Nesse sentido, a implementação dessas tecnologias requer a adoção de mecanismos rigorosos de monitoramento, de modo a assegurar a observância aos princípios do contraditório, da ampla defesa e da segurança jurídica.

Diante do exposto, observa-se que a regulação da desinformação e da proteção de dados, aliada ao uso ético e transparente da Inteligência Artificial, representa desafios da atualidade. O equilíbrio entre liberdade de expressão, proteção informacional e inovação tecnológica demanda soluções regulatórias que sejam, simultaneamente, eficazes e compatíveis com os direitos fundamentais assegurados em regimes democráticos.

A título de conclusão, o avanço na construção de arcabouços normativos que garantam um ambiente digital seguro, transparente e inclusivo revela-se imperativo para a preservação da esfera pública democrática na era da informação.

REFERÊNCIAS

BAPTISTA, Renata Ribeiro; AGUIAR, Julio Cesar de. Fake news, eleições e comportamento. **Revista Direito, Estado e Sociedade**, Ahead of Print, 2020. Disponível em: <https://revistades.jur.puc-rio.br/index.php/revistades/article/view/1320>. Acesso em: 9 fev. 2025.

BRASIL. Lei nº 13.709, de 14 de agosto de 2018. **Lei Geral de Proteção de Dados Pessoais (LGPD)**. Diário Oficial da União, Brasília, DF, 15 ago. 2018. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm. Acesso em: 9 fev. 2025.

BRASIL. Lei nº 7.716/1989. **Define os crimes resultantes de preconceito de raça ou de cor**. Diário Oficial da República Federativa do Brasil, Brasília, DF, 06 jan. 1989. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l7716.htm. Acesso em: 0 fev. 2025

BRASIL. Projeto de Lei nº 2.630, de 2020. **Lei Brasileira de Liberdade, Responsabilidade e Transparência na Internet**. Disponível em:

<https://www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=2252589>. Acesso em: 10 fev. 2025.

CASTELLS, Manuel. **O poder da comunicação**. Tradução: Rita Espanha. Lisboa: Fundação Calouste Gulbenkian, 2013.

DONEDA, Danilo; ARANHA, Estela. **O debate sobre o anonimato no caso Sleeping Giants Brasil. O Estadão**. 2020. Disponível em: <https://politica.estadao.com.br/blogs/fausto-macedo/o-debate-sobre-o--anonimato-no-caso-do-sleep-giants-brasil/>. Acesso em: 9 fev. 2025

EUROPEAN COMMISSION. **The Digital Services Act: ensuring a safe and accountable online environment**. Brussels, 2022. Disponível em: https://ec.europa.eu/commission/presscorner/detail/en/ip_22_2545. Acesso em: 10 fev. 2025.

GASPARETTO, Hígor Lameira; PEDROSO, Frederico Thaddeu; OLIVEIRA, Rafael Santos de. **Fake news**, discursos de ódio e ativismo digital: movimentos sociais de desmonetização, desafios jurídicos e reflexões sobre o case Sleeping Giants Brasil. *Revista Brasileira de Políticas Públicas*, Brasília, v. 12, n. 3, p. 179-198, 2022. Disponível em: <https://doi.org/10.5102/rbpp.v12i3.7960>. Acesso em: 9 fev. 2025.

HABERMAS, Jürgen. **A ética do discurso e a questão da verdade**. São Paulo: Martins Fontes, 2020.

KOETZ, Eduardo. Liberdade de expressão na era digital: desafios e perspectivas jurídicas. **Migalhas**, São Paulo, 11 jul. 2024. Disponível em: <https://www.migalhas.com.br/depeso/410997/liberdade-de-expressao-na-era-digital-desafios-e-perspectiva-juridica>. Acesso em: 9 fev. 2025.

ONU – ORGANIZAÇÃO DAS NAÇÕES UNIDAS. **Transformando Nosso Mundo: A Agenda 2030 para o Desenvolvimento Sustentável**. Assembleia Geral da ONU, 2015. Disponível em: <https://sustainabledevelopment.un.org/post2015/transformingourworld>. Acesso em: 9 fev. 2025.

SALINAS, Natasha Schmitt Caccia. Desafios de um modelo de regulação em rede para o combate às fake news. **Fábrica de Leis**, 30 maio 2023. Disponível em: <https://www.conjur.com.br/2023-mai-30/fabrica-leis-desafios-modelo-regulacao-rede-fake-news/>. Acesso em: 10 fev. 2025.

SARLET, Ingo Wolfgang. Liberdade de expressão e o problema da regulação do discurso do ódio nas mídias sociais. *Rei - Revista Estudos Institucionais*, v. 5, n. 3, p. 1207-1233, dez. 2019. DOI: <https://doi.org/10.21783/rei.v5i3.428>. Disponível em: <https://estudosinstitucionais.emnuvens.com.br/REI/article/view/428/443>. Acesso em: 9 fev. 2025.

SILVA, Rosane Leal da *et al.* Discursos de ódio em redes sociais: jurisprudência brasileira. **Revista Direito GV**, São Paulo, v. 7, n. 2, p. 445-468, dez. 2011. DOI: <https://doi.org/10.1590/S1808-24322011000200004>. Disponível em:

http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1808-24322011000200004&lng=en&nrm=iso. Acesso em: 9 fev. 2025.

SOUZA, Carlos Affonso. **Governança Digital e Proteção de Dados: Desafios e Perspectivas**. São Paulo: Revista dos Tribunais, 2021.

VASCONCELOS, Maria Inês. **O Marco Regulatório da Internet e a Filosofia do Direito: Liberdade de Expressão em Perspectiva**. Revista Direito e Tecnologia, ed. 06 dez. 2024.

Submetido em 05.10.2025

Aceito em 11.10.2025