

PUBLICAÇÃO

100

ISSN: 0101-9562

ISSN ELETRÔNICO: 2177-7055

SEQÜÊNCIA

Publicação do
Programa de Pós-Graduação
em Direito da UFSC

VOLUME 46 ■ ANO 2025

Estudos
jurídicos
e políticos



SEQÜÊNCIA – ESTUDOS JURÍDICOS E POLÍTICOS é uma publicação temática e de periodicidade quadrimestral, editada pelo Programa de Pós-Graduação Stricto Sensu em Direito da Universidade Federal de Santa Catarina – UFSC.

SEQÜÊNCIA – ESTUDOS JURÍDICOS E POLÍTICOS is a thematic publication, printed every four months, edited by the Program in law of the Federal University of Santa Catarina – UFSC.

Versão eletrônica: <http://www.periodicos.ufsc.br/index.php/sequencia>

Editora-Chefe: Norma Sueli Padilha

Editor Associado: José Sérgio da Silva Cristóvam

Editores Adjuntos: Priscilla Camargo Santos, Thanderson Pereira de Sousa

A publicação é indexada nas seguintes bases de dados e diretórios/

The Publication is indexed in the following databases and directories:

Base OJS

Base PKP

CCN (Catálogo Coletivo Nacional)

Dialnet

DOAJ (Directory of Open Access Journals)

EBSCOhost

Genamics Journalseek

Google Scholar

ICAP (Indexação Compartilhada de Artigos de Periódicos)

Latindex

LivRe!

ÖAW

OJS

PKP

Portal de Periódicos UFSC

Portal do SEER

ProQuest

SciELO

Scopus/Elsevier

Sherpa/Romeo

Sumarios.org

ULRICH'S

vLex

Ficha catalográfica

Seqüência: Estudos jurídicos e políticos. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Direito. n.1 (janeiro 1980)-.

Florianópolis: Fundação José Boiteux. 1980-.

Publicação contínua

Resumo em português e inglês

Versão impressa ISSN 0101-9562

Versão on-line ISSN 2177-7055

1. Ciência jurídica. 2. Teoria política. 3. Filosofia do direito. 4. Periódicos. I. Universidade Federal de Santa Catarina. Programa de Pós-graduação em Direito

CDU 34(05)

Catálogo na fonte por: João Oscar do Espírito Santo CRB 14/849

PUBLICAÇÃO		SEQÜÊNCIA	Estudos jurídicos e políticos

Redefinindo o alcance da atividade de moderação de conteúdo em plataformas digitais

Redefining the Scope of Content Moderation Activity on Digital Platforms

Vanice Valle¹

¹Universidade Federal de Goiás, Goiânia, Brasil.

RESUMO: A irradiação de plataformas digitais como arena de relações sociais diversas propõe o debate em torno da regulação destes ambientes, em contraste com a proposição de liberdade de expressão absoluta. O texto sustenta que essa discussão tem se travado a partir de uma visão reducionista do problema, que tem em conta tão-somente o texto da comunicação, desconsiderando outros elementos igualmente relevantes do ecossistema digital que podem interferir na construção da mensagem. A pesquisa aqui retratada se desenvolveu a partir de uma abordagem dialética; quanto ao seu objetivo, ela se desenvolve com a aplicação do método descritivo e exploratório. Aprofundando a compreensão da relação entre comunicação e o ambiente em que ela se desenvolve, o texto identifica os mecanismos hoje utilizados como reação às mensagens abusivas ou indesejadas, para depois propor componentes mínimos a serem considerados numa matriz regulatória que se pretenda efetiva na disciplina de comportamentos patológicos no exercício da expressão individual.

PALAVRAS-CHAVE: Moderação de conteúdo. Plataformas digitais. Liberdade de expressão. Ambiente da comunicação.

ABSTRACT: The spread of digital platforms as arenas for diverse social relations prompts debate surrounding the regulation of these environments, in contrast to the proposition of absolute freedom of expression. This text argues that this discussion has been conducted from a reductionist perspective, considering only the text of the post, disregarding other equally relevant elements of the digital ecosystem that can interfere in the construction of the message. The research presented here developed from a dialectical approach; regarding its objective, it



employs a descriptive and exploratory method. By deepening the understanding of the relationship between communication and the environment in which it develops, the text identifies the mechanisms currently used in response to abusive or unwanted messages, and then proposes minimum components to be considered in a regulatory framework intended to be effective in disciplining pathological behaviors in the exercise of individual expression.

KEYWORDS: Content moderation. Digital platforms. Freedom of expression. Communication ambience.

1. CONSIDERAÇÕES INICIAIS

Informa o módulo Tecnologia da Informação e Comunicação (TIC) da Pesquisa Nacional por Amostra de Domicílios (PNAD) Contínua publicado em 2024 (referente ao período 2023)(Brasil, [s/d]a), que 94,6% dos usuários brasileiros acessaram a internet para conversar por chamadas de voz ou vídeo. As outras finalidades mais relatadas foram: enviar ou receber mensagens de texto, voz ou imagens por aplicativos diferentes de e-mail (91,1%); assistir a vídeos, inclusive programas, séries e filmes (87,6%) e usar redes sociais (83,5%).

Essa massiva utilização de plataformas de comunicação do mundo digital, é disciplinada em nosso país, na sua maior porção até o momento, por uma matriz normativa que tem como destaque a Lei 12.965, de 23 de abril de 2014 (Marco Civil da Internet), à época precursora no tratamento de relações constituídas no ambiente digital –hoje, claramente insuficiente para a disciplina da miríade de relações que se estabelecem neste ecossistema. Mais recentemente, deu-se a edição da Lei 15.211, de 17 de setembro de 2025, denominada “Estatuto Digital da Criança e do Adolescente” – diploma evidentemente mais delimitado na sua esfera de cogitação, que veio à luz por força de um clamor social, curiosamente catalisado por um vídeo divulgado por influenciador digital¹.

¹ Para quem leia este texto no futuro, o movimento foi deflagrado pela publicação de um vídeo que cuidava daquilo que o então influenciador digital Felca denominou

Esse o contexto em que a regulação de mídias sociais ganha expressivo destaque, com a reiterada afirmação de que a construção deste arcabouço disciplinar deva ser pauta prioritária do Legislativo. No centro do debate, a par de potencial de violação a direitos fundamentais em geral, tem-se a questão atinente ao valor constitucionalmente protegido da liberdade de expressão, que alguns afirmam beneficiário de uma presunção absoluta de relevância social, qualquer que seja o conteúdo veiculado – do que resultaria verdadeira interdição à limitação de veiculação de conteúdos em ambiente digital. Em contraposição a essa tese, veicula-se o argumento de que não se tem na expressão livre um direito absoluto, e que a sua manifestação é de ter em conta os efeitos potenciais sobre o convívio social e o ambiente democrático.

Estabelecido o caráter central da disputa acerca dos contornos regulatórios passíveis de desenho para o funcionamento de plataformas de comunicação digital; tem-se a quase intuitiva identificação de dois pontos que estariam a merecer especial destaque na regulação: a definição de um ou mais modelos de moderação de conteúdo²; e o regramento do regime de responsabilidade civil a ser atribuído quando esta curadoria não se verifique de maneira adequada, observados os parâmetros aplicáveis, seja nos termos de uso da plataforma, seja no ordenamento jurídico vigente.

No campo da responsabilidade civil, tem-se a recente decisão do Supremo Tribunal Federal nos RE 1.037.396 (Tema 987 da Repercussão Geral) (Brasil, [s/d]b) e RE 1.057.258 (Tema 533 da Repercussão Geral) (Brasil, [s/d]c). É de se destacar, todavia, que mesmo

“adultização”, que constituiria na transposição de crianças e adolescentes para um ambiente de comportamentos e práticas próprias do mundo adulto como elemento de monetização de canais digitais.

² Diz-se um ou mais, porque remanesce o debate secundário em torno da viabilidade de uma única moldura regulatória, ou da superioridade técnica de uma solução fundada na ideia de segmentação por setores de operação de plataformas e mecanismos digitais.

esta matriz disciplinadora reporta os critérios de solução do conflito judicializado a uma tipologia de conteúdo veiculado, que se reputou impróprio a qualquer título. Reconduz-se o esquema de solução dos pontos críticos acima assinalados a uma abordagem reducionista do problema que se concentra na gestão do material de comunicação que vem à luz nas plataformas digitais.

É certo que a moderação de conteúdo não se constitui um fim em si mesma; mas é indiscutível que ela se poria a serviço da proteção de uma determinada concepção de liberdade de expressão a se aplicar no ambiente virtual, bem como de um padrão de uso compatível com a proteção a direitos fundamentais outros como a da proteção à privacidade e intimidade³.

O problema de pesquisa posto a este artigo, reside na indagação quanto à possibilidade e relevância de, num contexto de emergência de um quadro normativo orientado à moderação de conteúdo, identificar-se um mínimo regulatório que deva se ter por presente, de forma a prestigiar os valores jurídicos constitucionalmente consagrados como direcionadores da ação estatal. A hipótese é de que a concentração do debate público e parlamentar em torno exclusivamente de uma possível tipificação de conteúdos qualificados como indesejáveis traduz uma visão reducionista de um tema complexo, que não se revelará apta a efetivamente solucionar o problema público em discussão.

Constitui objeto deste artigo, a delimitação do que se entenda por moderação de conteúdo em ambiente digital; nisso se compreendendo não só os mecanismos possíveis para o seu exercício em caráter preventivo ou reativo, como também os pontos na cadeia de produção e veiculação deste mesmo material comunicativo em que essa atividade de controle deva se verificar. O objetivo é, com

³ Registre-se igualmente a pertinente preocupação com a relação entre conteúdo no ambiente digital e garantia de um ambiente democrático adequadamente informado e plural – mas esta questão transborda os limites propostos a este texto.

esse conhecimento aprofundado do fenômeno sob observação e dos desafios de sua curadoria, subsidiar não só a futura regulação em si, mas também previsíveis atualizações que hão de se verificar no futuro – afinal, não se presume que matéria tão volátil seja definitivamente equacionada com uma única investida regulatória.

Consideradas as limitações próprias a um artigo científico, o exercício argumentativo se dará tendo em conta especificamente a relação entre moderação de conteúdo e liberdade de expressão – eis que as demais violações potenciais a direitos fundamentais no ambiente digital frequentemente encontrarão nessa mesma atividade seu leito principal de veiculação. Destaque-se que o artigo não cogitará do delineamento de qual seja a concepção de liberdade de expressão que se possa extrair do texto constitucional. Esse debate, decerto dos mais relevantes nos dias de hoje, não caberia nos estreitos limites do artigo acadêmico, e a rigor, não precisa ser enfrentado para o aprofundamento do tema de uma arquitetura desejável à atividade de moderação de conteúdo.

O artigo se desenvolverá, observada a seguinte sequência de conteúdo. No item 1, as considerações de ordem metodológica. O item 2 destaca que o exercício de expressão no universo do digital tem também no ambiente em que ele se dá, um relevante componente da comunicação e mensagem. Disso decorre que o foco exclusivamente no conteúdo se revela insuficiente a uma regulação adequada. O item 3, municiado pela compreensão dos desafios do novo ambiente a ser regulado, percorre as técnicas e mecanismos hoje utilizados para fins de moderação de conteúdo, num esforço de verificação da aptidão para dar resposta ao desafio regulatório posto. Numa perspectiva propositiva, o item 4 se ocupa de indicar componentes a serem considerados numa matriz regulatória de moderação de conteúdo. Finalmente, o item 5 de conclusão reforça a importância de um conhecimento mais amplo do contexto em que o problema se põe, como pressuposto ao desenvolvimento de uma regulação efetiva.

A pesquisa aqui retratada se desenvolveu a partir de uma abordagem dialética; quanto ao seu objetivo, ela se desenvolve com a aplicação do método descritivo e exploratório.

Apresenta-se como resultado, a demonstração da relevância de que a regulação de plataformas digitais de comunicação não deixe de dispor sobre: 1) reconhecimento da relevância do comunicador no ambiente digital como elemento identificador do regime de responsabilidade; 2) clareza nos processos *by design* e por provocação, de gestão de ocorrências na veiculação de conteúdo; e 3) a importância de um ranqueamento de infrações como métrica de confiabilidade da comunicação veiculada na plataforma.

2. RESGATE DA RELEVÂNCIA DO MEIO NA REGULAÇÃO DAS PLATAFORMAS DIGITAIS

Constitui discurso corrente aquele segundo o qual as novas tecnologias, notadamente plataformas de comunicação em geral e mecanismos de inteligência artificial, estão a merecer urgentemente, regulação adequada. O desafio posto ao atendimento desta demanda está na identificação de modelos de regência de relações jurídicas neste novo ambiente digital, que se revelem aptos a produzir os efeitos de ordenação eficaz da prática social desejados, sem prejuízo a direitos fundamentais implicados no desenvolvimento da comunicação; em especial, aquele à livre expressão. Militam contra os modelos tradicionais de regulação jurídica já experimentados, dentre outros fatores, a velocidade e ubiquidade próprias do ambiente digital.

Neste contexto, identificar o que seja típico do ambiente digital – seja no modo de interação, no controle da arena de comunicação; seja nos agentes envolvidos no processo de troca e influência mútua – se apresenta como passo primeiro à exploração de alternativas regulatórias. É preciso conhecer o que o ambiente digital apresenta como problema público distinto de outros já anteriormente parametrizados,

como TV, rádio, jornalismo etc. É só a partir desta compreensão que se poderá construir uma matriz de regulação adequada.

Costuma-se indicar como elemento diferenciador do quadro social presente – portanto, orientador dos termos em que a matriz regulatória seria de se construir para responder aos desafios do momento – a circunstância de que a disseminação de comunicações de toda ordem, no ambiente digital se dá de maneira quase instantânea, praticamente sem limitações geográficas, alcançando um número exponencial de destinatários⁴. Esse caráter quase síncrono das interações se apresentaria como um desafio à regulação, eis que incrementa em muito o potencial de dano decorrente de mensagens tidas por socialmente indesejáveis, e torna um desafio o propósito de erigir um mecanismo formal de reação que se movimente à mesma velocidade. O apontamento, todavia, pode se revelar, se não enganoso, quando menos simplificador. Afinal, é certo que o referido atributo se aplica à comunicação que hoje se dá no universo digital – mas isso se verifica, a rigor, como culminância de uma curva de transformação em relação aos meios anteriormente disponibilizados para esse mesmo exercício de intercâmbio e expressão de ideias. Significa dizer que o fenômeno, a rigor, não é inteiramente novo – talvez a inovação se situe tão-somente no plano da intensidade com que ele se manifesta.

A percepção de instantaneidade na circulação de informações hoje associada às plataformas digitais é fenômeno já havido no passado, quando da incorporação, por exemplo, do rádio; veículo que permitia a oitiva síncrona de pessoas diretamente envolvidas no acontecimento noticiado, ou ainda discursos e pronunciamentos de autoridades⁵. A

⁴ Diz-se exponencial eis que por vezes, a comunicação empreendida no ambiente digital pode encontrar eco em outros agentes, causando a chamada viralização, não necessariamente pelo alcance do comunicador original, mas pela reprodução de terceiros.

⁵ O primeiro discurso transmitido pelo rádio, do então Rei George V, alcançou súditos por todo o mundo – Canada, Nova Zelândia, Austrália, África do Sul, Gibraltar, e Irlanda (British Heritage Travel, 2023).

TV de outro lado proporcionou em 1969, que 600 milhões de pessoas acompanhassem a chegada do homem à lua (Memória Globo, 2021) – comunicação essa que ultrapassou barreiras geográficas, e suplantou qualquer estimativa inicial de número de possíveis assistentes. Em ambas as ocasiões, deu-se o estranhamento quanto ao alcance da mensagem – antes tido como impensável – proporcionado pelos meios de comunicação disponíveis em cada qual dos acontecimentos. Não é no atributo da quase sincronicidade portanto, que se terá a particularização das novas plataformas de comunicação ofertadas no ambiente digital.

Ponto de diferenciação que se pode associar especificamente ao ambiente digital – e com implicações relevantes no universo da regulação jurídica – é a dissociação que ele permite, entre o emitente original da mensagem e os alcançados por ela, com o uso de mecanismos como a repostagem, a colaboração, o *remix*, o *repost* e outras funcionalidades assemelhadas. Tais ferramentas oportunizam uma hipertextualidade, cujo (des)valor é de ser objeto de maior reflexão. Afinal, dela podem resultar efeitos como a dissipação da autoridade do comunicador original ou a ocultação desse mesmo autor; o esgarçamento de mecanismos de curadoria dos termos da mensagem, que circula descontroladamente alcançando outros ambientes menos supervisionados que seu ambiente original de procedência. Finalmente, mas não menos importante, é de se considerar ainda a potencial distorção, dolosa ou culposa do conteúdo inicial da mensagem. Práticas como a descontextualização e a agregação de interpretações que se somam à comunicação primeva, obscurecem a autoria, e podem subverter a comunicação como originalmente intencionada.

Esse quadro de potencial dissociação entre o comunicador original e o recipiente da mensagem se agrava com a disseminação dos mecanismos de inteligência artificial generativa, que permitem a criação inteiramente fantasiosa de avatares ou versões digitais de pessoas públicas, emitindo mensagens inteiramente falsas. Mecanismo ainda mais sutil são intervenções na comunicação original com o uso de IA

generativa; intervenções essas que podem passar despercebidas, ao menos do público em geral, menos municiado de informação técnica necessária a identificar a manipulação.

Outros atributos são associados ao ambiente digital, a saber, o potencial de circulação – eis que as plataformas têm expressivo padrão de interconexão, permitindo, quando não estimulando, que a comunicação gerada no ambiente “A” seja instantaneamente reproduzida no ambiente “B”. Agregadores de comunicação digital viabilizam que uma mesma postagem seja, a um só tempo, veiculada em vários ambientes, potencializando o número de alcançados pela mensagem. A par disso, os mecanismos de priorização na distribuição aplicados pelos algoritmos não são conhecidos pelos usuários, circunstância que pode determinar o encaminhamento da comunicação a públicos que não os originalmente cogitados pelo emissor; público esse que, eventualmente por não se valer dos mesmos signos, pode receber uma mensagem substantivamente distinta daquela que se desejava.

Observe-se que os atributos típicos do ambiente digital acima descritos são decorrência não do conteúdo da comunicação em si, mas da arquitetura dos meios em que elas circulam. Hipertextualidade e interconexão são efeitos de uma estrutura de plataformas digitais que contempla esse tipo de circulação da mensagem. Esta constatação permite afirmar que o combate a comportamentos patológicos no uso de plataformas digitais é atividade que não pode se limitar à curadoria do que se contenha em textos ou imagens numa postagem original; mas precisa abarcar uma visão mais ampla do fenômeno de comunicação que se verifica nestes ambientes.

O debate público posto no cenário nacional tem sido capitaneado pela ideia de que a curadoria de conteúdo de mensagens exercida por aparatos digitais como mecanismos de inteligência artificial, ou ainda por agentes humanos, seja o principal desafio regulatório. Afinal, é nesse campo que se pode identificar a patologia da hora do ambiente digital: as chamadas *fake news*, assim entendidas como novas mensagens que contenham informação incorreta ou falsa, e que não declinam

a existência da incorreção ou falsidade (Wang, 2020)⁶. Afirmar-se a inadequação dos mecanismos estatais disponíveis para a oferta de resposta tempestivo e proporcional aos riscos postos por essa distorção comunicacional (Cunha Filho, de Carvalho e Carvalho, 2022), e com isso a agenda regulatória passa a se concentrar no controle do conteúdo da mensagem, atividade que se supõe possa ser desenvolvida de modo satisfatório e *by design* por mecanismos de alinhamento que operem por trás daqueles de processamento de linguagem natural e/ou de imagens.

A concentração das atenções nesse componente do problema tem contribuído para uma polarização entre defensores de uma regulação que tem em conta principalmente o conteúdo da mensagem e seu potencial bloqueio; e aqueles que se classificam como paladinos da garantia de liberdade de expressão, e que vislumbram grave risco de censura neste exercício de um juízo valorativo quanto à comunicação. Já aqui se identifica uma distorção na compreensão do problema público, eis que não há uma oposição insuperável entre as duas posições. É possível projetar modelos regulatórios que harmonizem comunicação e liberdade de expressão – mas para tanto, é preciso compreender que no ambiente digital, a comunicação não se dá exclusivamente através do texto ou da imagem.

Importante enriquecer o quadro de análise dos desafios postos à comunicação e à informação veraz, com a incorporação do elemento negligenciado, a saber, os efeitos que o meio possa determinar sobre a mensagem. Afinal, se existe essa relação de influência do veículo no qual circula a comunicação, e a percepção daqueles alcançados pela mensagem em relação ao seu conteúdo; uma regulação que verse só sobre este último pode se revelar insuficiente à garantia dos resultados pretendidos.

Para incorporar essa dimensão no quadro geral do desafio regulatório, útil será a formulação de McLuhan (2006, p. 23), que

⁶ Para uma categorização das possíveis manifestações de *fake news* que contribui para a compreensão do fenômeno a ser combatido, consulte-se Tandor Jr., Lee e Ling (2018).

já ao final da década de 60 advertia que “o meio é a mensagem, porque é o meio que configura e controla a proporção e a forma das ações e associações humanas”. Significa dizer que o meio pode adir à comunicação original, nuances que interferem na mensagem – tudo isso com ou sem a ciência e concordância do emissor inicial. A ordem de apresentação de postagens, com ou sem reiteração de uma mesma tipologia de conteúdo dele ou de outro autor, pode determinar no destinatário da comunicação reações de reforço da mensagem, aprofundando vieses de confirmação; ou pode ainda, caso se apresente pontos de vista distintos, incentivar a formulação pelo receptor, de um juízo crítico ou ainda o exercício de uma saudável dialética.

É ainda do mesmo autor (McLuhan, 2006, p. 22) a advertência de que “a ‘mensagem’ de qualquer meio ou tecnologia é a mudança de escala, cadência ou padrão que esse meio introduz nas coisas humanas”. Esta proposição reforça o ponto já apresentado segundo o qual o desafio regulatório esteja umbilicalmente relacionado aos atributos do meio (ambiente digital) agora oferecido à comunicação.

Primeiro reflexo da proposição de McLuhan (2006) na compreensão do problema da moderação de conteúdo em plataformas digitais, é a mudança de escala na percepção quanto ao conteúdo – fenômeno que se tem por alinhado com a heurística da disponibilidade. Algoritmos de distribuição conferem predominância a determinados tipos de comunicações, a partir de critérios que combinam a *persona* digital do usuário, com parâmetros de conveniência da plataforma digital enquanto agente econômico. Isso gera para o usuário uma percepção de escala da relevância de opiniões ou efeitos de uma determinada comunicação que pode não ser consentânea com a realidade – mas fato é que a interferência do meio conforma a recepção da mensagem pelo usuário. Na sua interface com a heurística da disponibilidade, essa mudança de escala pode alterar, por exemplo, a percepção de risco (Sunstein, 2012) – e como se sabe, a suposição de uma ameaça pode determinar reações extremas em indivíduos ou grupos sociais.

Extraí-se ainda da referência de McLuhan (2006) a advertência da potencial mudança de cadência que meio e tecnologia podem oferecer. Esse fenômeno, quando se tem em conta o universo do digital e suas plataformas de comunicação, diz respeito não só à aceleração ao extremo da circulação em si da mensagem, mas também de um suposto imperativo de reação daqueles que são por ela alcançados, seja pela produção de uma “resposta” à comunicação inicial, seja pelo seu rápido compartilhamento. Essa mudança de cadência agrega ao desafio regulatório, não só a construção de mecanismos de supressão tempestiva da mensagem indesejável; mas também a necessidade de incorporar ferramentas que harmonizem essa cadência diferenciada do digital, com os conhecidos cânones da ampla defesa e do devido processo legal – opostos sempre à providência de contenção da comunicação pelos defensores de uma liberdade de expressão ampliada.

Finalmente, no campo da interferência em relação a padrões estabelecidos nas relações humanas, tem-se um componente que merece destaque, a saber, a mudança no critério de atribuição de autoridade do comunicador, que passa a compreender (para muitos) a proporção de sua presença neste mesmo ambiente digital, traduzida normalmente no número de seguidores, *likes* ou visualizações. Assim, tem-se por diluída a importância de outros signos de autoridade previamente estabelecidos, como escolaridade, experiência pessoal ou profissional etc., em favor desse indicador numérico, cuja origem, autenticidade e causa, a rigor não se conhece. Cabe então refletir se a regulação de conteúdo é de se aplicar segundo os mesmos critérios àquele que tem estabelecida essa “autoridade digital”, na hoje tão propalada figura do influenciador; e àquele que simplesmente se comunica com sua limitada comunidade de amigos e familiares.

Mais ainda; é indiscutível que os meios de comunicação oferecidos neste admirável mundo novo – e o ambiente digital parece ser o ponto alto desta curva de crescimento – permitem uma mobilização de sentidos humanos, e ainda de outros componentes que, na arena da economia da atenção, não só conseguem captar o olhar do usuário,

mas direcionar a sua percepção quanto ao conteúdo da mensagem. Cores, conteúdo de imagens, caráter estático ou móvel; todos esses são elementos sobejamente conhecidos por profissionais de *marketing* digital, que deles se valem na disputa pela atenção. Essa mobilização de sentidos, por sua vez, é construída a partir de critérios algorítmicos que, processando e classificando as comunicações que circulam no ambiente digital, estabelecem relação de precedência ou hierarquia entre comunicações considerado o perfil do usuário; e ainda criam liames de conexão ou similitude onde elas talvez não existam. Aqui também se identifica uma manifestação da reconfiguração da autoridade, que passa a ser traduzida por um posicionamento da comunicação expresso algorítmicamente (Pasquale, 2015) – e pior, por um algoritmo cujos termos não se conhece.

Não se está, com estas considerações, sustentando uma posição conducente à compreensão de que se tenha na tarefa de moderação de conteúdo uma missão impossível. O que se pretende é evidenciar que limitar a deliberação pública sobre a matéria ao debate em torno da tipificação de conteúdos indesejáveis; da tentativa de desenho de um procedimento de curadoria de texto ou imagem; ou ainda de um regime de responsabilidade civil pretensamente reparador do dano já materializado resultará num quadro regulatório apequenado, inapto a responder aos desafios que o problema propõe.

3. MECANISMOS DE MODERAÇÃO DE COMUNICAÇÕES NO AMBIENTE DIGITAL: ETAPAS PASSADAS E QUADRO ATUAL

O tema da moderação de conteúdo no ambiente digital não se constitui, a rigor, uma novidade – mas é matéria posta desde os primórdios de utilização da Internet. Ainda que se tenha em conta a chamada Internet 1.0, de formato estático, *read only*, ali já se tinha, ainda que em muito menor escala, a figura do produtor de conteúdo, problematizando a questão relacionada a eventual responsabilidade

por danos associados ao material disponibilizado. Em contraposição aos receios atinentes a eventual causação de danos por mensagens veiculadas no ambiente digital, punha-se o temor de que uma regulação muito restritiva pudesse sufocar uma atividade que poderia permitir em tese, não só ganhos econômicos, mas também um incremento significativo na circulação de ideias e conhecimento.

A opção à época, envolveu um desenho que apostava predominantemente em eximir de responsabilidade o produtor de conteúdo, dando mais espaço não só para o desenvolvimento da florescente tecnologia segundo uma orientação econômica (Zuboff, 2020) que tem em conta um mercado de comportamentos futuros, mas também para o exercício livre da expressão humana. É o caso, por exemplo, da opção estadunidense pela redação conferida à *Section 230* do *Communications Decency Act* (Lirio do Valle, 2023), que eximia de responsabilidade os provedores de plataformas digitais de comunicação. É possível vislumbrar alguma influência dessa modelagem mesmo no quadro legal brasileiro, com a Lei 12.965, de 23 de abril de 2014 (Marco Civil da Internet), cujo art. 19, lido segundo uma aproximação mais afeita à literalidade, conduziu durante muito tempo à compreensão de que a responsabilidade do produtor de conteúdo se restringia às hipóteses de desatendimento a ordem judicial – percepção só alterada com as recentes decisões do Supremo Tribunal Federal nos Temas 987 e 533 da Repercussão Geral.

A eclosão da Internet 2.0, em que a produção e veiculação de conteúdo é compartilhada com todos os usuários, reveste o problema de mais complexidade, com milhões e milhões de pessoas concorrendo para a formatação de mensagens em várias mídias; mensagens essas que, embora possam corresponder à visão de mundo de seu emissor, não traduzem necessariamente um discurso tido por aceitável na arena pública, especialmente tendo em conta que esta última, hoje, pode compreender todo o planeta, com suas incontáveis nuances de cultura. Mais do que isso, a estrutura de curadoria de comunicações, que é própria por exemplo, dos órgãos de imprensa tradicionais, decerto

não pode ser estabelecida em relação a cada qual dos comunicadores atuando no espaço digital. Assim, mecanismos de proteção quanto à adequação do discurso, *fact checking* e outros, comuns aos veículos tradicionais de informação e notícia se perderam, no mundo dos influenciadores e das “tias do zapp”. O desafio é múltiplo, potencializado não só pela velocidade com que as comunicações circulam, mas também pela variedade de formatos de que elas se revestem. A par disso, o significado de palavras, símbolos e imagens pode variar de local a local, ou ainda em função do tempo; tudo segundo a cultura ali estabelecida⁷. O resultado é que um conteúdo reputado normal no ambiente “a” pode se mostrar ofensivo na localidade “b”, o que acrescenta camadas de complexidade à eventual moldura regulatória. Não se pode igualmente desprezar a influência do tempo nas comunicações, como se verifica nos episódios de “cancelamento” de figuras públicas, onde o esquadramento de mensagens veiculadas no passado remoto segundo padrões de aceitabilidade então vigentes conduz ou reforça o juízo de reprovação social.

A contraposição do imperativo de garantia da liberdade de expressão como impedimento absoluto a iniciativas de moderação de conteúdo, é estratégia que perde força, quando se tem ainda nesse campo, uma sofisticação de discurso que acolhe uma teoria democrática de liberdade de expressão (Fiss, 2022). Nesta perspectiva, renuncia-se a um predomínio de uma visão individualista, centrada na potência e relevância da autoexpressão; para valorizar a expressão como um componente relevante de um debate público qualificado, que permita, sempre que cabível, autodeterminação coletiva formulada a partir de informações confiáveis e de um diálogo respeitoso. Da adoção de uma teoria democrática quanto aos limites à liberdade de expressão,

⁷ Recentemente, a série “Adolescência”, veiculada no canal de *streaming* Netflix tematizava, em cena reveladora, como *emojis* utilizados em redes sociais tinham significados absolutamente distintos para gerações diferentes que travassem contato com esse mesmo símbolo.

decorre o reconhecimento de que o simples manejo de mecanismos limitados de responsabilização, como nos modelos cunhados no início da década de 90, se revelaria insuficiente a prevenir os danos potenciais ao ambiente e debate público associados a mensagens reputadas inadequadas por qualquer razão.

Uma outra vertente relacionada ao tema da moderação de conteúdo é o reconhecimento de que o ecossistema digital se estende para muito além das plataformas que exibem as comunicações a serem moderadas. A par dos diversos atores situados na camada de veiculação do conteúdo (mídias sociais, serviços de veiculação de vídeos, mercados *online*, fóruns de discussão, aplicativos de economia compartilhada etc.), tem-se ainda aqueles que integram a camada de distribuição, despontando nesta os aplicativos de busca; e ainda os situados no nível de oferta de infraestrutura como os provedores, os registradores de domínio, autoridades certificadoras etc.

Compreendida a dimensão do ecossistema digital, tem-se por claro que a atividade de moderação de conteúdo envolve inicialmente a expressão de decisões formuladas por agentes privados em relação a terceiros – antes de ser uma deliberação decorrente de uma manifestação judicial ou da ação corretiva de uma instância administrativa de qualquer ordem. O problema reside em que esta decisão privada pode gerar consequências relevantes para seus destinatários diretos (emissor da mensagem) ou indiretos (usuário). Nestes termos, o modelo de formulação da moderação de conteúdo pelo agente privado, que acontece também antes da veiculação, é de ser objeto de igual cuidado, eis que também nessa instância é possível identificar-se a violação não só à liberdade de expressão, mas ainda de outros direitos fundamentais do emissor da comunicação relacionados em especial à sua defesa, mas também daqueles que com ele interagem.

No cenário corrente, limitadas são no Brasil as alternativas de coibição de comunicações reputadas inadequadas ou danosas a qualquer título – moderação de conteúdo na sua hipótese mais gravosa. A primeira ferramenta é a imputação de responsabilidade civil, suscetível

de endereçamento em face do autor da comunicação. Consigne-se que mesmo este mecanismo, oriundo ainda da primeira fase da reflexão sobre moderação de conteúdo, é objeto de debate judicial junto ao STF no Tema 987 da Repercussão Geral (Brasil, 2017) no que toca à suposta limitação à responsabilidade civil dos provedores de serviços por força do art. 19 do Marco Civil da Internet⁸.

A par da eventual imputação do dever de reparação de danos, o quadro atual disponibiliza – seja por força do ordenamento vigente, seja pelas cláusulas de serviço das plataformas –, em relação ao autor da comunicação, as seguintes providências: 1) suspensão da postagem em si; 2) suspensão da conta do usuário; e 3) impedimento definitivo de veiculação de comunicações do usuário de plataformas digitais, em perfil próprio ou de terceiros, no território nacional. Mais recentemente, considerado o potencial de monetização de determinadas plataformas de comunicação, também a suspensão deste tipo de contrapartida tem sido utilizada como mecanismo de coerção a publicações reputadas inadequadas (Brasil, 2022); assim como a vedação ao impulsionamento das mesmas comunicações. Verifica-se, portanto, uma ampliação paulatina dos mecanismos específicos de resposta à postagem indesejável, incorporando providências que tenham por objeto o emissor, cujo padrão de comportamento se reputa tóxico.

⁸ Tenha-se em conta uma certa hesitação da Corte na delimitação da análise a ser empreendida. Isto se diz pelo potencial descompasso entre o título atribuído ao Tema 987 (Discussão sobre a constitucionalidade do art. 19 da Lei n. 12.965/2014 (Marco Civil da Internet) que determina a necessidade de prévia e específica ordem judicial de exclusão de conteúdo para a responsabilização civil de provedor de internet, websites e gestores de aplicativos de redes sociais por danos decorrentes de atos ilícitos praticados por terceiros), e a descrição que a mesma página institucional oferece em relação ao mesmo tema (Recurso extraordinário em que se discute, à luz dos arts. 5º, incs. II, IV, IX, XIV e XXXVI, e 220, caput, §§ 1º e 2º, da Constituição da República, a constitucionalidade do art. 19 da Lei n. 12.965/2014 (Marco Civil da Internet) que impõe condição para a responsabilização civil de provedor de internet, websites e gestores de aplicativos de redes sociais por danos decorrentes de atos ilícitos de terceiros).

É certo que uma moldura regulatória que avance sobre o terreno do modo segundo o qual empresas privadas empreendem à moderação de conteúdo pode se revelar um desafio. Afinal, é de se reconhecer a estes agentes econômicos liberdade para alinhar seus serviços de oferta de plataformas digitais aos objetivos e valores da empresa privada, tudo nos termos da garantia que é também constitucional, da livre iniciativa e da livre concorrência (art. 170, *caput* e inciso IV da CF). De outro lado, sua atividade compreende a oferta de serviços e informações ao público em geral, o que permite ao Estado *lato sensu* a regulação, quando menos, por força de outro princípio regedor da ordem econômica, que é a proteção ao consumidor; para não mencionar outras hipóteses que tem o signo de valoração constitucional.

Destaque-se, como referência de solução normativa recente para o problema público sob exame, o *Digital Services Act* (DSA) (European Union 2022), adotado pela União Europeia, que opta por uma regulação que tem por principal eixo orientador o reconhecimento do caráter de atividade de risco inerente aos provedores de serviços digitais. A partir desta premissa, o DSA constrói uma arquitetura regulatória que muda o foco da responsabilização *ex post* pela veiculação de conteúdo indesejável, para uma ideia de *accountability by design*, onde as plataformas passam a ter deveres legais de diligência em relação a seus juízos de veiculação ou de interdição de conteúdo.

O tema, não obstante suas dificuldades, já encontra visitaç o em terra *brasilis*. A Resolu  o 23.732, de 27 de fevereiro de 2024 (Brasil, 2024), que altera a anterior Resolu  o 23.610, de 18 de dezembro de 2019 (Brasil, 2019) – ambas do Tribunal Superior Eleitoral, incorpora j a a dissemina  o dos mecanismos de intelig ncia artificial generativa, para atribuir deveres espec ficos de informa  o quanto ao uso deste tipo de recurso, favorecendo que o eleitor possa formular o devido ju zo cr tico quanto   credibilidade do afirmado ou exibido em imagens que tem o concurso destes aparatos digitais. Tem-se nessa estipula  o, comando de modera  o de conte do que

tem nos produtos da comunicação o seu destinatário primário – mas também nas plataformas veiculadoras das mensagens, um dever de desenvolvimento de *due diligence*, com a exigência ainda de planejamento e execução de ações preventivas. Preceituações como essa indicam uma sensibilidade já externada, ao menos pela Justiça Eleitoral, à caracterização da veiculação de comunicações no mundo digital como atividade de risco.

A literatura atual explora a identificação e valoração desse elemento de risco a partir da perspectiva do usuário (Branley e Covey, 2018; Ahern, Sauer e Thacker, 2015): quais as ameaças postas a ele, seja pela sua presença em ambiente digital que se revele tóxico, seja pela interação danosa com as plataformas digitais sob análise. É de se incorporar a uma agenda de pesquisa futura, se não há outras perspectivas de análise quanto à existência de riscos vinculados à veiculação de conteúdos falsos, inadequados ou indesejáveis.

4. COMPONENTES A SEREM CONSIDERADOS NUMA MATRIZ REGULATÓRIA DE MODERAÇÃO DE CONTEÚDO

Revisitado o problema público posto, cumpre agora contribuir para a construção de caminhos de superação, com o indicativo de componentes que se reputa devam ser incorporados a uma matriz regulatória eficaz, que guarde relação de congruência com os termos em que as comunicações se dão no ambiente digital.

Parece relevante reconhecer numa iniciativa regulatória, que existe diferenciação entre a presença ordinária, dos cidadãos comuns no ambiente virtual, e aquele dos operadores que tem relevância sistêmica. Isso é de se aplicar em relação a provedores de serviços, mas igualmente no que toca aos comunicadores ou (para usar a expressão corrente) influenciadores – sem os quais, ainda que o ofertante de serviços tenha uma carteira grande de acessos, um resultado nocivo significativo parece improvável.

A lógica da gestão de riscos, já incorporada em arcabouços regulatórios de países outros, seria de se estender essa mesma diferenciação quanto a deveres especiais de cautela e proteção também ao emissor da comunicação, que sendo – para usar a expressão da área – “grande no digital”, deve empreender à sua própria análise quanto ao potencial danoso, seja em comunicações de sua própria autoria, seja naquelas em que se “surfa na onda” de postagem viral de terceiros. Observe-se ainda que produtores de conteúdo com presença expressiva no ambiente digital tem potencial de monetização evidentemente mais alto, o que permitira superar o argumento (sempre frágil) dos custos atinentes a atividades como a verificação de veracidade.

Segundo componente relevante numa matriz regulatória vocacionada à regulação de conteúdo diz respeito à clareza nos processos desenvolvidos e aplicados pela organização – seja naqueles estatuídos *by design* no desenvolvimento e nos critérios ordinários de operação da plataforma, de prevenção à divulgação de conteúdo tóxico; seja naqueles relacionados ao tratamento conferido pela mesma plataforma à provocação atinente à existência de comunicação indesejável.

Observe-se que o apontamento aqui formulado diz respeito à visibilidade ampliada não dos mecanismos internos de operação em si do sistema – informação que, ainda que se tenha como possível, no mais das vezes será inútil ao cidadão comum, e mesmo às pequenas organizações com presença no universo digital. O que se está destacando como importante na matriz regulatória é a exigência de conhecimento formal da estrutura de moderação de conteúdo (se automatizada ou analógica, com ou sem supervisão humana; em que termos se desenvolve; qual o tipo de aparato do provedor de serviços digitais que assume a moderação e as reclamações; quais os componentes da comunicação considerados etc.⁹). Destaque-se que essa mesma transparência é de se estender em

⁹ Reitera-se a observação antes formulada de que a comunicação transcende os limites da linguagem, e compreende contexto. Nestes termos, conhecer se a moderação *by*

relação ao tratamento conferido, seja ao protesto formulado pelo autor ou replicador pela supressão do conteúdo; seja à reclamação do alcançados pelo conteúdo impróprio que pretende sua supressão.

A clareza em relação a esse tipo de informação pode, de um lado, funcionar como elemento inibidor da tentativa de postagem tóxica – afinal, conhecidos os mecanismos de controle, a iniciativa pode parecer àquele que dela cogita, revestida de baixa probabilidade de êxito. De outro lado, o conhecimento dos mecanismos disponíveis para correção ou subtração de conteúdos indesejáveis pode facilitar a movimentação por aquele que se sinta afetado. O resultado seria o empoderamento de terceiros atingidos pela comunicação, mas também a ampliação, ainda que por via indireta, dos agentes de moderação de conteúdo – universo que passaria a contemplar a parcela da sociedade informada quanto às regras do jogo.

Finalmente, mas não menos importante, é de se ter em conta que a toxicidade da comunicação falsa está intimamente ligada a um elemento (real ou imaginário) de confiabilidade do veiculador desta mesma informação. Nestes termos, pode ser interessante o estímulo à divulgação de um ranqueamento de *strikes* (providência de censura a conteúdo impróprio desenvolvida pela plataforma) – medida que permitiria ao receptor do conteúdo um mínimo de verificação em relação à autoridade de quem comunica.

Não se pretende ter nestes atributos um elenco exaustivo, até porque a resiliência do ambiente digital é conhecida, o que evidencia de plano o caráter sempre provisório de iniciativas regulatórias. O objetivo é emancipar a visão quanto à moderação de conteúdo, para que se possa compreender o processo de comunicação em toda a sua complexidade.

design compreende um espectro mais amplo de elementos desta mesma comunicação é um componente relevante, seja na avaliação do resultado de reprobado empreendido em concreto, seja naquela da habilidade em abstrato do provedor de serviços digitais de operar preventivamente de forma eficaz.

5. CONCLUSÃO

A ubiquidade do digital evidencia cada vez mais que as fórmulas tradicionais de comando e sanção são insuficientes ao enfrentamento dos riscos associados a esse admirável mundo novo. Nem por isso, podemos abdicar do Direito e sua normatividade como ferramenta de ordenação do convívio, e de proteção a valores socialmente relevantes.

O esforço é de ser no sentido de atualizar as matrizes de incidência do Direito sobre relações sociais e jurídicas que se forjam a partir de outras premissas que não as ditadas até há pouco pelo mundo do analógico. Isso porque a fórmula por vezes repetida de que “o que não pode ser feito na vida real, no mundo real, não pode ser feito no mundo virtual” (Agência Brasil, 2023), é lamentavelmente simplificadora, e apequena os esforços de regulação que se vem desenvolvendo no ambiente legislativo.

Compreender o mundo virtual, envolve uma aproximação cognitiva maior quanto aos atributos que este mesmo ambiente traz para o desenvolvimento de atividades cujos modelos de execução conhecemos no plano do analógico – o que não significa que a simples transposição seja a solução adequada. É certo que os exemplos de usos questionáveis de plataformas digitais se multiplicam na experiência da vida, mas matrizes regulatórias com efetivo condão de prevenir padrões de comportamento indesejáveis devem partir de uma compreensão ampla do contexto em que a atividade se desenvolve.

É verdade que Arthur C. Clark já advertia que “qualquer tecnologia suficientemente avançada é indistinguível da magia”; e isso pode funcionar como um desincentivo a aprofundamento maior no plano do mágico. Mas a frase acima, é preciso entender, é só uma alegoria; o que precisamos evitar a todo custo é o engano de que problemas com a complexidade trazida pelo mundo digital se resolvam simplesmente com a varinha de condão – ou a caneta.

REFERÊNCIAS

- AHERN, Nancy R.; SAUER, Penny; THACKER, Paige. Risky behaviors and social networking sites: how is YouTube influencing our youth? *Journal of Psychosocial Nursing and Mental Health Services*, v. 53, n. 10, p. 25-29, 2015.
- BRANLEY, Dawn Beverley; COVEY, Judith. Risky behavior via social media: the role of reasoned and social reactive pathways. *Computers in Human Behavior*, v. 78, p. 183-191, 2018.
- BRASIL. Câmara dos Deputados. Projeto de Lei nº 2.630/2020. Institui a Lei Brasileira de Liberdade, Responsabilidade e Transparência na Internet. Disponível em: <https://www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=2256735>. Acesso em: 27 jun. 2024.
- BRASIL. Instituto Brasileiro de Geografia e Estatística. *Acesso à internet e à televisão e posse de telefone móvel celular para uso pessoal 2023*. Agência IBGE Notícias, [s.d.].a. Disponível em: https://agenciadenoticias.ibge.gov.br/media/com_mediaibge/arquivos/b433a668d9b22364c742c3fab5c8f227.pdf. Acesso em: 22 out. 2025.
- BRASIL. Supremo Tribunal Federal. Recurso Extraordinário 1.037.398. Rel. Min. Dias Toffoli, [s.d.].b. Sem acórdão à época da elaboração do texto; ajuizado em 29 mar. 2017. Tema 987 da Repercussão Geral. Disponível em: <https://portal.stf.jus.br/jurisprudenciaRepercussao/verAndamento-Processo.asp?incidente=5160549&numeroProcesso=1037396&classeProcesso=RE&numeroTema=987>. Acesso em: 27 jun. 2024.
- BRASIL. Supremo Tribunal Federal. Recurso Extraordinário 1.057.258. Rel. Min. Luiz Fux, [s.d.].c. Sem acórdão à época da elaboração do texto; ajuizado em 20 out. 2011 como ARE 660861. Tema 533 da Repercussão Geral. Disponível em: <https://portal.stf.jus.br/jurisprudenciaRepercussao/pesquisarProcesso.asp?txtTituloTema=533>. Acesso em: 27 jun. 2024.
- BRASIL. Tribunal Superior Eleitoral. Resolução nº 23.610, de 18 de dezembro de 2019. Dispõe sobre a propaganda eleitoral. Disponível em: <https://www.tse.jus.br/legislacao/compilada/res/2019/resolucao-no-23-610-de-18-de-dezembro-de-2019>. Acesso em: 27 jun. 2024.

BRASIL. Tribunal Superior Eleitoral. Ação de Investigação Judicial Eleitoral n. 0601522-38.2022.6.00.0000. Rel. Min. Benedito Gonçalves, julg. em 20 out. 2022.

BRASIL. Tribunal Superior Eleitoral. Resolução nº 23.732, de 27 de fevereiro de 2024. Altera a Res.-TSE nº 23.610, de 18 de dezembro de 2019, dispondo sobre a propaganda eleitoral. Disponível em: <https://www.tse.jus.br/legislacao/compilada/res/2024/resolucao-no-23-732-de-27-de-fevereiro-de-2024>. Acesso em: 27 jun. 2024.

BRITISH HERITAGE TRAVEL STAFF. The first Royal Christmas Message, by King George, in 1932, was written by Rudyard Kipling. *British Heritage Travel*. Disponível em: <https://britishheritage.com/history/first-royal-christmas-message>. Acesso em: 30 set. 2024.

CUNHA FILHO, Marcio; DE CARVALHO, Pedro Feitosa Araújo; CARVALHO, Sofia. Fake news: definições, tipologias e a insuficiência das respostas estatais. *Revista de Estudos Empíricos em Direito*, v. 9, p. 1-35, 2022.

EUROPEAN UNION. Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and amending Directive 2000/31/EC (Digital Services Act). Disponível em: https://eur-lex.europa.eu/legal-content/EN/TXT/?toc=OJ%3AL%3A2022%3A277%3ATOC&uri=uriserv%3AO-J.L._2022.277.01.0001.01.ENG. Acesso em: 30 set. 2024.

EUROPEAN UNION. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending various EU regulations (Artificial Intelligence Act). Disponível em: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ%3AL_202401689. Acesso em: 7 ago. 2024.

FISS, Owen M. *A ironia da liberdade de expressão: Estado, regulação e diversidade na esfera pública*. Tradução e prefácio de Gustavo Binenbojm; Caio Mário da Silva Pereira Neto. 2. ed. Rio de Janeiro: FGV Editora, 2022.

LEIS DE CLARKE. *Wikipedia*, [s.d.]. Disponível em: https://pt.wikipedia.org/wiki/Leis_de_Clarke. Acesso em: 30 out. 2025.

LÍRIO DO VALLE, Vanice Regina. Regulação de plataformas digitais: uma agenda propositiva à luz dos leading cases de judicial review no Brasil e nos EUA. *A&C – Revista de Direito Administrativo & Constitucional*, Belo Horizonte, v. 23, n. 94, p. 133-157, 2023. DOI: 10.21056/aec.v23i94.1805.

Disponível em: <https://www.revistaaec.com/index.php/revistaaec/article/view/1805>. Acesso em: 26 jun. 2024.

MCLUHAN, Marshall. *Os meios de comunicação como extensões do homem (Understanding Media)*. Tradução de Décio Pignatari. São Paulo: Editora Cultrix, 2006.

MEMÓRIA GLOBO. Chegada do homem à Lua. *O Globo*, 29 out. 2021. Disponível em: <https://memoriaglobo.globo.com/jornalismo/coberturas/chegada-do-homem-a-lua/noticia/chegada-do-homem-a-lua.ghtml>. Acesso em: 26 jun. 2024.

PASQUALE, Frank. *The black box society: the secret algorithms that control money and information*. Cambridge, Massachusetts; London, England: Harvard University Press, 2015.

PEDUZZI, Pedro. Moraes: regras do mundo real devem prevalecer também nas redes sociais. *Agência Brasil*, 18 abr. 2023. Disponível em: <https://agenciabrasil.ebc.com.br/justica/noticia/2023-04/moraes-regras-do-mundo-real-devem-prevalecer-tambem-nas-redes-sociais>. Acesso em: 18 abr. 2023.

SUNSTEIN, Cass. Para além do princípio da precaução. *RDA – Revista de Direito Administrativo*, Rio de Janeiro, v. 259, p. 11-71, jan./abr. 2012.

TANDOC JR., Edson C.; LIM, Zheng Wei; LING, Richard. Defining “fake news”: a typology of scholarly definitions. *Digital Journalism*, v. 6, n. 2, p. 137-153, 2018.

WANG, Chih-Chien. Fake news and related concepts: definitions and recent research development. *Contemporary Management Research*, v. 16, n. 3, p. 145-174, 2020.

ZUBOFF, Shoshana. *A era do capitalismo de vigilância: a luta por um futuro humano na nova fronteira do poder*. Tradução de George Schlesinger. Rio de Janeiro: Intrínseca, 2020.

VANICE VALLE

A autora é Visiting Fellow no Human Rights Program da Harvard Law School, pós-doutora pela EBAPE/FGV e doutora em Direito

pela Universidade Gama Filho. Atuou como Procuradora do Município do Rio de Janeiro e exerceu funções de gestão pública em secretarias municipais e estaduais de administração. É professora permanente do PPGDP da UFG, com extensa produção acadêmica e atuação em cursos de graduação e pós-graduação. Também produz conteúdo digital na área jurídica, incluindo o Programa GPS Acadêmico, além de pesquisa em direito público, direitos fundamentais e jurisdição constitucional.

Endereço Profissional: Universidade Estácio de Sá, Programa de Pós-Graduação em Direito. Rua São José, nº 35, 15º andar. Centro.

ORCID ID: <https://orcid.org/0000-0002-7857-6486>

E-MAIL: vanicevalle@gmail.com

Recebido em: 31/10/2025

Aceito em: 08/12/2025

Editor responsável:

Dr. Thanderson Pereira de Sousa

Dra. Norma Sueli Padilha

Disponibilidade dos Dados

A autoria declara que todos os dados utilizados na pesquisa encontram-se disponíveis em repositório público, em conformidade com as práticas de ciência aberta. A Revista Sequência estimula o compartilhamento de dados de pesquisa que assegurem a transparência, a reprodutibilidade e a verificação dos resultados publicados, respeitando, entretanto, os princípios éticos aplicáveis. Assim, não é exigida a divulgação de informações que permitam a identificação de sujeitos de pesquisa ou comprometam sua privacidade. O compartilhamento de dados deve, portanto, priorizar a integridade científica e a proteção de dados sensíveis, garantindo a publicização dos resultados sem exposição indevida de participantes.



Este trabalho está licenciado sob uma licença Creative Commons Attribution 4.0 International License.

Autores e autoras cedem à Revista Sequência direitos exclusivos de primeira publicação, ficando o trabalho licenciado sob a Creative Commons Attribution 4.0 International License. A licença autoriza que terceiros remixem, adaptem e ou criem a partir do trabalho publicado, indicando o crédito ao trabalho original e sua publicação inicial. Os autores têm permissão para assumir contratos adicionais em separado, com distribuição não exclusiva da versão publicada na Revista Sequência, indicando, de todo modo, a autoria e publicação inicial neste periódico.