



UNIVERSIDADE DE BRASÍLIA (UnB)
FACULDADE DE DIREITO (FD)
CURSO DE GRADUAÇÃO EM DIREITO

OTÁVIO HENRIQUE MAYRINK RESENDE

**CONSTRUINDO NOVOS CONSENSOS PARA PROTEGER A
LIBERDADE DE EXPRESSÃO NA INTERNET:**

Uma análise do PL 2.630/2020 a partir de uma *perspectiva procedimental* para regular as grandes plataformas digitais.

BRASÍLIA

2021

OTÁVIO HENRIQUE MAYRINK RESENDE

**CONSTRUINDO NOVOS CONSENSOS PARA PROTEGER A
LIBERDADE DE EXPRESSÃO NA INTERNET:**

Uma análise do PL 2.630/2020 a partir de uma *perspectiva
procedimental* para regular as grandes plataformas digitais.

Trabalho de Conclusão de Curso
apresentado como requisito parcial à
obtenção do grau de Bacharel em Direito
pela Universidade de Brasília.

Orientador: Prof. Dr. Alexandre Kehrig
Veronese Aguiar

BRASÍLIA

2021

FICHA CATALOGRÁFICA:

[RESENDE, Otávio].

[Título: CONSTRUINDO NOVOS CONSENSOS PARA PROTEGER A LIBERDADE DE EXPRESSÃO NA INTERNET: uma análise do PL 2.630/20 a partir de uma perspectiva procedimental para regular as grandes plataformas digitais] / Otávio H. Mayrink Resende. — 2021. 108 f.

Trabalho de Conclusão de Curso (Bacharelado em Direito) – Universidade de Brasília (UnB), Brasília, 2021.

1. [primeira entrada de assunto]. 2. [segunda entrada de assunto]. 3. [terceira entrada de assunto]. I. Título.

CDD [número da CDD].

REFERÊNCIA BIBLIOGRÁFICA:

RESENDE, Otávio H. Mayrink. **Construindo novos consensos para proteger a liberdade de expressão na internet: uma análise do PL 2.630/20 a partir de uma perspectiva procedimental para regular as grandes plataformas digitais.** Monografia de Final de Curso (Graduação em Direito), Faculdade de Direito, Universidade de Brasília, DF, 108f, 2021.

OTÁVIO HENRIQUE MAYRINK RESENDE

CONSTRUINDO NOVOS CONSENSOS PARA PROTEGER A LIBERDADE DE EXPRESSÃO NA INTERNET: Uma análise do PL 2.630/2020 a partir de uma *perspectiva procedimental* para regular as grandes plataformas digitais.

Trabalho de Conclusão de Curso apresentado como requisito parcial à obtenção do grau de Bacharel em Direito pela Universidade de Brasília.

Aprovado em: 05 de novembro de 2021.

BANCA EXAMINADORA

Prof. Dr. Alexandre Kehrig Veronese Aguiar (Orientador)
Universidade de Brasília

Prof.^a Dra. Ana de Oliveira Frazão (Avaliadora)
Universidade de Brasília

Prof.^a Dra. Ana Cláudia Farranha Santana (Avaliadora)
Universidade de Brasília

Prof. Ms. João Victor Archegas (Avaliador)
Harvard Law School

AGRADECIMENTOS

Agradeço a Deus, que me mostrou que nada faz sentido sem amor.

Agradeço aos meus pais, maiores apoiadores, que sonham junto comigo e dão o exemplo de integridade e trabalho com a própria vida. Agradeço também à minha irmã, pelas inúmeras demonstrações de carinho e afeto que me fazem lembrar das coisas realmente importantes da vida.

Agradeço ao meu orientador, o professor Alexandre Veronese, pelo apoio incondicional, paciência, e valiosas contribuições que foram fundamentais para a construção desse trabalho.

Agradeço ao querido mestre Airton Seelaender, referência ética, profissional e acadêmica cujo exemplo de excelência e rigor intelectual me marcaram profundamente e com quem aprendi a pensar criticamente o Direito.

Agradeço aos tantos amigos e colegas com quem tive a alegria de estudar, conviver e trabalhar no Laboratório de Pesquisa em Políticas Públicas e Internet (LAPIN). Mais do que um centro de pesquisa, o LAPIN foi um laboratório de aprendizados de vida, e um projeto conjunto que me enche de orgulho.

Agradeço à querida professora Ana Farranha, pelo carinho, responsabilidade, e ajuda que foi fundamental para manter de pé esse projeto, que hoje alça voos próprios.

Agradeço ao professor Vallisney de Souza Oliveira, de quem testemunhei a paixão pelos livros e seu poder transformador, e de quem guardo tenras lembranças como aluno e voluntário do projeto Habeas Liber.

Agradeço à comunidade acadêmica da Faculdade de Direito da Universidade de Brasília (UnB), que juntamente com meus excelentes estágios no Ministério da Ciência, Tecnologia, Inovações e Comunicações (MCTIC), no escritório Beltrão Advocacia, no Supremo Tribunal Federal (STF) e no Conselho Administrativo de Defesa Econômica (CADE), despertaram em mim o interesse pelo Direito.

Enfim, minha mais profunda gratidão por todas as amizades que construí, pelos mestres que tive, pelos conhecimentos que adquiri, e pelas ricas experiências que vivi nesta etapa da minha vida.

“All men are constrained by means external to them to ends equally external. The further the technical mechanism develops that allows us to escape natural necessity, the more we are subjected to artificial technical necessities.” (ELLUL, 1954, p. 429)

RESUMO

As grandes plataformas digitais moldam, estruturam e mediam a forma como as pessoas interagem, emitem opiniões e se relacionam umas com as outras na internet. Estas empresas também controlam o modo como os conteúdos são produzidos, distribuídos e consumidos no espaço virtual, estabelecendo para isso normas privadas (os “termos de serviço” ou “padrões da comunidade”) que são aplicadas aos usuários por meio da moderação de conteúdo. Face a essa dinâmica de “autorregulação corporativa” da governança de conteúdo online, diversos governos e parlamentos ao redor do mundo têm procurado formas de regular esses serviços, em especial as redes sociais. Há, contudo, um impasse aparentemente insuperável nesse debate. De um lado, argumenta-se que as plataformas digitais têm moderado conteúdo de forma excessiva, comprometendo direitos e liberdades individuais. Do outro, defende-se que as empresas não têm moderado conteúdo suficientemente, sendo necessário combater práticas prejudiciais como a desinformação e os discursos de ódio, impondo restrições à liberdade de expressão quando necessário. Iniciativas legais assentadas nessas visões trazem frequentemente respostas regulatórias desproporcionais, ou capazes de ampliar o risco de violações à liberdade de expressão. Para superar esse impasse, propõe-se uma abordagem regulatória focada nos *procedimentos decisórios* da governança de conteúdo online e não propriamente em juízos substanciais sobre o que deve ou não ser permitido na internet. Essa *perspectiva procedimental* propugna, em suma, o reconhecimento de novos direitos e deveres na relação dos usuários com as plataformas digitais (ex.: dever de cuidado, obrigação de defender a integridade dos usuários, dever de transparência etc.), bem como a adoção de mecanismos de devido processo na moderação de conteúdo (ex.: direito ao recurso, direito à notificação e a decisões fundamentadas etc.) e a criação de uma autoridade com independência técnica e decisória para avaliar casos concretos, e garantir o cumprimento de suas decisões. Partindo dessa premissa, este trabalho analisou as soluções contidas no texto do Projeto de Lei nº 2.630/2020 aprovado pelo Senado Federal em junho de 2020. Concluiu-se que, conquanto a proposta legislativa preveja novos direitos e garantias processuais aos usuários nos procedimentos de moderação, foi possível identificar problemas na redação legislativa que podem dar azo a arbitrariedades, bem como discrepâncias no desenho regulatório que propõe um modelo ineficaz de “autorregulação regulada”, e não apresenta garantias reais de proteção aos direitos fundamentais e a bens jurídicos caros ao interesse público.

Palavras-chave: Liberdade de expressão; Moderação de conteúdo; Regulação; Devido processo; PL 2.630/2020; Constitucionalismo digital; Responsabilização de intermediários; Governança de plataformas.

ABSTRACT

Large digital platforms shape, structure and mediate the way people interact, emit opinions, and relate to each other on the internet. These companies also control how content is produced, distributed, and consumed in the virtual space, establishing private standards for this (the “terms of service” or “community standards”) that are applied to users through content moderation. Faced with this dynamic of “corporate self-regulation” of online content governance, several governments and parliaments around the world have been looking for ways to regulate these services, especially social networks. There is, however, a seemingly insurmountable impasse in this debate. On the one hand, it is argued that digital platforms have excessively moderated content, compromising individual rights and freedoms. On the other hand, it is argued that companies do not have enough moderate content, making it necessary to combat harmful practices such as disinformation and hate speech, imposing restrictions on freedom of expression when necessary. Legal initiatives based on these views often bring regulatory responses that are disproportionate, or capable of increasing the risk of violations of freedom of expression. To overcome this impasse, we propose a regulatory approach focused on decision-making procedures for online content governance rather than on substantive judgments about what should or should not be allowed on the internet. This procedural perspective advocates, in short, the recognition of new rights and duties in the relationship of users with digital platforms (eg, the duty of care, obligation to defend the integrity of users, the duty of transparency, etc.), as well as adoption due process mechanisms in content moderation (e.g., right to appeal, right to notification and reasoned decisions, etc.) and the creation of an authority with technical and decision-making independence to assess concrete cases, and ensure compliance with its decisions. Based on this premise, this work analyzed the solutions contained in Bill No. 2,630/2020, also known as "PL das Fake News", approved by the Brazilian Federal Senate in June 2020. It was concluded that, although the legislative proposal provides for new rights and procedural guarantees for users in moderation procedures, it was possible to identify problems in the legislative wording that can give rise to arbitrariness, as well as discrepancies in the regulatory design that proposes an ineffective model of "regulated self-regulation" and does not present real guarantees of protection of rights fundamental and public interest.

Keywords: Freedom of speech; Content moderation; Regulation; Due process; PL 2630/2020; Digital Constitutionalism; Intermediary liability; Platform governance.

LISTA DE ABREVIATURAS E SIGLAS

ADI	Ação Direta de Inconstitucionalidade
Anatel	Agência Nacional de Telecomunicações
CADE	Conselho Administrativo de Defesa Econômica
CCTCI	Comissão de Ciência e Tecnologia, Comunicação e Informática
CETIC	Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação
CFRB	Constituição Federal da República do Brasil
CDA	<i>Communications Decency Act</i>
CDR	Coalizão Direitos na Rede
CEJ	Corte Europeia de Justiça
CGI	Comitê Gestor da Internet no Brasil
CIDH	Comissão Interamericana de Direitos Humanos
CNCPC	Conselho Nacional dos Chefes de Polícia Civil
CNJ	Conselho Nacional de Justiça
CNMP	Conselho Nacional do Ministério Público
Conar	Conselho Nacional de Autorregulamentação Publicitária
CTRI	Conselho de Transparência e Responsabilidade na Internet
DMCA	<i>Digital Millennium Copyright Act</i>
DPF	Departamento de Política Federal
DUDH	Declaração Universal de Direitos Humanos
EFF	<i>Electronic Frontier Foundation</i>
EUA	Estados Unidos da América
FOB	<i>Facebook Oversight Board</i>
GTNET	Grupo de Trabalho de Aperfeiçoamento da Legislação Brasileira na Internet
IDEC	Instituto Brasileiro de Defesa do Consumidor
LAPIN	Laboratório de Políticas Públicas e Internet
NetzDG	<i>Netzwerkdurchsetzungsgesetz</i>
MCI	Marco Civil da Internet
MCTIC	Ministério da Ciência, Tecnologia, Inovações e Comunicações
MI	Mandado de Injunção
MP	Medida Provisória
ONU	Organização das Nações Unidas
PIDCP	Pacto Internacional sobre Direitos Civis e Políticos

PEC	Proposta de Emenda à Constituição
PL	Projeto de Lei
STF	Supremo Tribunal Federal
TJ-SP	Tribunal de Justiça de São Paulo
UE	União Europeia
UK	<i>United Kindgdom</i>

SUMÁRIO

1	INTRODUÇÃO	11
1.1	NOTAS METODOLÓGICAS	16
2	UMA PERSPECTIVA PROCEDIMENTAL PARA REGULAR A LIBERDADE DE EXPRESSÃO NA INTERNET	17
2.1	A LIBERDADE DE EXPRESSÃO COMO DIREITO FUNDAMENTAL	17
2.1.1	<i>A liberdade de expressão no direito brasileiro</i>	19
2.1.1.1	<i>A responsabilização civil por conteúdo gerado por terceiros</i>	24
2.1.2	<i>A liberdade de expressão no direito internacional</i>	27
2.2	OS NOVOS CONTORNOS REGULATÓRIOS DA LIBERDADE DE EXPRESSÃO NA ERA DIGITAL	29
2.3	O TRIÂNGULO DE BALKIN: UM MODELO PLURALISTA DE REGULAÇÃO DA LIBERDADE DE EXPRESSÃO	32
2.3.1	<i>Problemas</i>	35
2.3.1.1	Censura prévia e colateral	35
2.3.1.2	Uma governança de conteúdo pouco transparente	37
2.3.1.3	Vigilância e manipulação	39
2.3.2	<i>Soluções propostas</i>	40
2.3.2.1	Regulações pró-competição	40
2.3.2.2	As responsabilidades da infraestrutura da internet	42
2.3.2.3	As obrigações de confiabilidade, ética e profissionalismo	42
2.3.2.4	As obrigações legais do devido processo	44
2.3.2.5	As obrigações legais dos fiduciários da informação	46
2.4	AS ERAS REGULATÓRIAS DA GOVERNANÇA PRIVADA DE CONTEÚDO ONLINE	49
2.4.1	<i>A Era dos Direitos</i>	51
2.4.2	<i>A Era da Saúde Pública</i>	54
2.4.3	<i>A Era do Processo</i>	55
2.4.3.1	<i>Soluções propostas</i>	56
2.5	UMA ABORDAGEM PROCEDIMENTAL PARA REGULAR AS GRANDES PLATAFORMAS DE CONTEÚDO: CONSENSOS, PROPOSTAS E DESAFIOS	57
2.5.1	<i>O reconhecimento de novos direitos e deveres</i>	58
2.5.2	<i>Entre uma “regulação autoritária” e uma “autorregulação corporativa”</i>	59
2.5.3	<i>O “devido processo” na moderação de conteúdo</i>	60
2.5.3.1	Direito à informação e dever de transparência	61
2.5.3.2	Direito à notificação e à decisão fundamentada	62
2.5.3.3	Direito de defesa e de apelação	63
2.5.4	<i>Um organismo de controle independente</i>	64
2.5.5	<i>Desafios e limitações</i>	65
3	O PROJETO DE LEI Nº 2630/2020	68
3.1	CONTEXTUALIZAÇÃO	68
3.2	DISPOSIÇÕES PRELIMINARES	69

3.3	OS PROCEDIMENTOS DE MODERAÇÃO	71
3.3.1	<i>Direito à informação e à mecanismos de recurso e devido processo.....</i>	71
3.3.2	<i>Direito à notificação, à decisão fundamentada e ao contraditório</i>	72
3.3.3	<i>Direito ao recurso</i>	73
3.3.4	<i>Direito à reparação por danos decorrentes de decisão equivocada.....</i>	73
3.3.5	<i>Direito a prazo de defesa maior no caso de deepfakes.....</i>	74
3.3.6	<i>Direito de resposta proporcional ao agravo.....</i>	74
3.4	O CONSELHO DE TRANSPARÊNCIA E REPONSABILIDADE NA INTERNET (CTRI).....	75
3.5	UMA INSTITUIÇÃO DE AUTORREGULAÇÃO REGULADA	76
3.6	AS SANÇÕES	77
3.7	DISPOSIÇÕES FINAIS RELEVANTES	77
4	ANÁLISE DO PL 2.630/2020 À LUZ DA PERSPECTIVA PROCEDIMENTAL DE REGULAÇÃO DAS GRANDES PLATAFORMAS DIGITAIS	78
4.1	CONSIDERAÇÕES INICIAIS	78
4.2	DISPOSIÇÕES PRELIMINARES	79
4.2.1	<i>Escopo</i>	79
4.2.2	<i>Objetivos</i>	80
4.3	AS REGRAS DE MODERAÇÃO.....	82
4.3.1	<i>Transparência e o direito à informação</i>	82
4.3.2	<i>O direito ao devido processo e ao recurso.....</i>	83
4.3.3	<i>O direito à notificação e à decisão fundamentada.....</i>	84
4.3.4	<i>Outros casos: reparação de danos, direito a resposta e deepfakes.....</i>	85
4.4	AS INOVAÇÕES INSTITUCIONAIS	85
4.4.1	<i>O Conselho de Transparência e Responsabilidade na Internet (CTRI)...</i>	85
4.4.2	<i>Uma instituição de autorregulação regulada</i>	87
4.5	SANÇÕES	88
4.6	CONCLUSÃO	89
5	CONSIDERAÇÕES FINAIS.....	90
6	REFERÊNCIAS	95

1 INTRODUÇÃO

Em 6 de setembro de 2021, o presidente Jair Bolsonaro editou a Medida Provisória nº 1068, que alterava significativamente o Marco Civil da Internet¹, e estabelecia novas regras para a moderação² de conteúdo nas redes sociais. A norma condicionava a exclusão, suspensão ou bloqueio de contas e conteúdo à apresentação de “justa causa e motivação”, nas hipóteses expressamente indicadas pelo Poder Executivo (BRASIL, 2021a).

Também estavam previstos novos direitos e salvaguardas processuais para os usuários e garantias à informação relacionadas à governança de conteúdo nas plataformas digitais. O descumprimento das novas regras poderia acarretar sanções cíveis, criminais ou administrativas, além de multas, suspensão das atividades e até a proibição do serviço no país (BRASIL, 2021a).

Em nota, o Palácio do Planalto afirmou que as redes sociais “passaram a exercer um papel fundamental na intermediação de relações pessoais e profissionais” da população e que os “meios eletrônicos de comunicação se tornaram um relevante instrumento para a manifestação de ideias e opiniões” (BRASIL, 2021b). O comunicado ainda justificava a urgência e relevância da decisão ao afirmar que:

“[...] a remoção arbitrária e imotivada de contas, perfis e conteúdos por provedores de redes sociais, além de prejudicar o debate público de ideias e o exercício da cidadania, resulta em um quadro de violação em massa de direitos e garantias fundamentais como a liberdade de expressão e o exercício do contraditório e da ampla defesa” (BRASIL, 2021b).

¹ A Lei nº 12.965 de 23 de abril de 2014, também conhecida como Marco Civil da Internet, estabelece princípios, garantias, direitos e deveres para o uso da Internet no Brasil (BRASIL, 2014).

² Neste trabalho, os termos “moderação”, “curadoria” ou mesmo “regulação” de conteúdo são utilizados para descrever o mesmo processo pelo qual os provedores de serviços online (ex.: mídias sociais, mecanismos de busca etc.) promovem a análise e a triagem do conteúdo gerado por seus usuários, determinando se estes se adequam ou não às regras definidas em seus termos de serviço, ou mesmo às normas de determinada localidade ou jurisdição. Essas regras definem não apenas “o que pode e o que não pode ser publicado”, mas também “qual conteúdo será visto primeiro e quais não serão”. Em caso de violações, as plataformas podem aplicar sanções temporárias ou permanentes como a remoção total ou parcial do conteúdo publicado, a imposição de restrições ao acesso do usuário à plataforma e suas funcionalidades, a desmonetização de um perfil ou conteúdo, a limitação do alcance das publicações etc. (ROBERTS, 2017; ONU, 2018, p. 3; KLONICK, 2018, p. 1601; IDEC, 2021).

A medida foi anunciada às vésperas dos atos convocados para o Dia da Independência, em relação aos quais havia o temor de incitações antidemocráticas. O texto provocou forte reação de políticos, acadêmicos e organizações da sociedade civil que consideraram a medida desproporcional e apontaram³ motivações político-eleitorais na providência (MONTEIRO, 2021).

Dentre as principais críticas levantadas estava a de que a publicação interferiria indevidamente na autonomia privada e na liberdade contratual das empresas, impedindo, na prática, que estas moderassem conteúdo com base em suas próprias regras. Sob a justificativa de “reforçar direitos e garantias de usuários”, o Poder Executivo prescreveu um rol taxativo de hipóteses para a moderação de conteúdo, sem incluir ali, entre outras coisas, a prática da desinformação ou discursos de ódio em geral. Nestes casos, as vítimas precisariam recorrer à via judicial para a remoção de conteúdo, caracterizando um verdadeiro “salvo-conduto” a tais práticas nocivas, que, em geral, demandam respostas rápidas para que seus danos sejam contidos (SOUZA, 2021; SARLET, 2021; SAFERNET, 2021; DOUEK, 2019, p. 10).

Foram propostas⁴ seis Ações Diretas de Inconstitucionalidade (ADIs) e um mandado de segurança foi impetrado contra a norma. Contudo, pouco mais de uma semana após sua publicação, decisões quase simultâneas do Supremo Tribunal Federal (STF) e da mesa diretora do Congresso Nacional puseram fim à providência, afastando sua eficácia (AMORIM *et al*, 2021).

A controvérsia ao redor da MP 1068 mostra como o tema da governança de conteúdo *online* tem dividido a opinião pública. De um lado, estão aqueles que argumentam que as plataformas digitais têm moderado conteúdo de forma excessiva, silenciando ideias e opiniões, em uma clara violação da liberdade de expressão. Do outro, há um grupo que entende que as empresas não têm moderado conteúdo suficientemente, sendo necessário combater práticas daninhas como a

³ O presidente Jair Bolsonaro já teve diversos vídeos removidos de plataformas digitais como YouTube, Twitter, Facebook, e Instagram sob a justificativa de propagar desinformação no contexto da pandemia de COVID-19 capaz de “causar danos reais às pessoas”. O presidente também alega que tem sido frequentemente “cerceado” em sua liberdade de expressão pelas redes sociais. Além disso, vários de seus apoiadores influentes tiveram suas contas suspensas, ou desmonetizadas por violarem as regras das plataformas (AMADO, 2021; UCHINAKA, 2021; VIEIRA, 2021; SENRA, 2020).

⁴ [ADI 6996](#), [ADI 6995](#), [ADI 6994](#), [ADI 6993](#), [ADI 6992](#), [ADI 6991](#) e o [MS 38207](#).

desinformação e os discursos de ódio, impondo restrições à liberdade de expressão quando necessário (BOWERS & ZITTRAIN, 2020, p. 7).

Este impasse, aparentemente insuperável, tem desafiado legisladores e reguladores em todo o planeta. Como resposta, diversas propostas têm emergido, nem sempre capazes de endereçar adequadamente os problemas.

Dentre as iniciativas regulatórias mais emblemáticas, está a *Netzwerkdurchsetzungsgesetz* (também conhecida como Lei de Vigilância de Redes ou “NetzDG”), lei alemã aprovada em 2017 que obriga as plataformas a remover em até 24 horas conteúdo considerado “manifestamente ilegal” pela legislação local. Caso não cumpram as regras, os provedores de aplicação de internet⁵ poderão sofrer duras sanções financeiras (AL SUR, 2020, p. 29).

Se por um lado, forçar as empresas a monitorar ostensivamente o conteúdo de seus usuários, a fim de identificar e reprimir supostas ilegalidades pode parecer uma estratégia eficiente para conter ilícitos na internet, por outro, esse método cria incentivos para a prática da censura. Caso os provedores de aplicações não removam rapidamente conteúdo supostamente ilegal, poderão sofrer punições severas. Em uma dinâmica na qual a plataforma atua como procurador, juiz, júri e executor, sem que usuários afetados tenham a oportunidade de serem ouvidos ou contem com as proteções do devido processo⁶, incentivos desta natureza tendem a fomentar ainda mais arbitrariedades (BALKIN, 2018, p. 2010-2018).

No Brasil, o Poder Legislativo tem discutido um marco normativo sobre esse tema no âmbito do Projeto de Lei 2.630/2020, popularmente conhecido como “PL das Fake News” (BRASIL, 2021c). A proposta legislativa em questão foi aprovada no Senado Federal em junho de 2020, e agora tramita na Câmara dos Deputados.

⁵ De acordo com Ceroy (2014), o termo “provedor de aplicação de internet” designa “qualquer empresa, organização ou pessoa natural que, de forma profissional ou amadora, forneça um conjunto de funcionalidades que podem ser acessadas por meio de um terminal conectado à internet, não importando se os objetivos são econômicos.”. Por essa definição, podem ser considerados provedores de aplicação os serviços de e-mail (ex.: Outlook, Gmail etc.), as redes sociais (ex.: Instagram, Facebook, Twitter, YouTube, etc.), os serviços de mensageria privada (ex.: WhatsApp, Telegram, etc.) os geradores de conteúdo (ex.: g1, UOL, etc.), os serviços de comércio eletrônico (ex.: OLX, Mercado Livre, etc.) e até sites de pequenas empresas podem ser incluídos nessa definição (ARRUDA, 2021).

⁶ Neste trabalho monográfico, considera-se que o “devido processo” na moderação de conteúdo incorpora, no mínimo, “a garantia da igualdade de tratamento, que as decisões sejam motivadas, que seja possível o exercício de uma defesa efetiva por parte dos usuários afetados, a possibilidade de apelar das decisões e que os procedimentos tenham uma duração razoável” (PALLERO, 2020, p. 18)

O presente trabalho explora uma possível saída para esse impasse, concentrando-se nos *procedimentos* decisórios da governança de conteúdo *online*. Isto é, ao invés de discutir o que deve ou não ser permitido na internet, busca-se construir um amplo consenso sobre *como* decisões sobre conteúdo devem ser tomadas. Propõe-se, portanto, deslocar o foco regulatório da definição de juízos *substanciais* sobre conteúdo, para disciplinar o *processo* de tomada dessas decisões (BOWERS & ZITTRAIN, 2020, p. 7; MOSTERT, 2020, p. 14).

Essa abordagem regulatória, denominada “perspectiva procedimental” propugna, entre outras coisas, o reconhecimento de novos direitos e deveres na relação dos usuários com as plataformas digitais (ex.: dever de cuidado, obrigação de defender a integridade dos usuários, dever de transparência etc.), e a adoção de mecanismos do devido processo na moderação de conteúdo (ex.: direito ao recurso, direito à notificação e a decisões fundamentadas etc.). Também se defende a necessidade de uma autoridade com garantias de independência, autonomia técnica e decisória, com o distanciamento necessário para avaliar casos concretos, e garantir o cumprimento de suas decisões (MONTEIRO *et al*, 2021, p. 28-30; PALLERO *et al*, 2020; BERGMAYER, 2018, p. 30; ARHEGAS *et al*, 2021, p. 21; FRAZÃO, 2021).

A partir dessa perspectiva, a *pergunta de pesquisa* que se quer responder é de que modo o PL 2.630/20, no texto aprovado pelo Senado Federal em junho de 2020, dispõe sobre garantias processuais para regular a governança privada de conteúdo na internet, tomando como prisma de análise a perspectiva procedimental.

Busca-se, desse modo, como *objetivo geral*, descobrir se as soluções regulatórias processuais dispostas pelo legislador brasileiro endereçam adequadamente os problemas e disfuncionalidades da moderação de conteúdo na era digital, partindo da abordagem regulatória supracitada.

Estabelece-se, para tanto, os seguintes *objetivos específicos*: (i) identificar os elementos jurídicos e conceituais que caracterizam a perspectiva procedimental para regular as grandes plataformas digitais; (ii) arrolar os dispositivos do PL 2.630/20 que dizem respeito a essa perspectiva regulatória; e (iii) analisar e discutir as provisões legislativas previamente relatadas à luz da abordagem regulatória delimitada.

A *hipótese* central é de que o PL 2.630/20 incorpora importantes avanços na regulação da governança de conteúdo online, como a previsão de direitos e garantias

dos usuários na moderação de conteúdo, mas traz problemas de técnica legislativa e institui um modelo de autorregulação⁷ regulada que, na forma como foi desenhado, pode dar azo a arbitrariedades e visões policiaiscas na moderação de conteúdo, indo de encontro a recomendações da *perspectiva procedimental*.

A *metodologia* mais apropriada para alcançar os objetivos delineados é a revisão bibliográfica e documental dos principais marcos teóricos relativos a este debate, especificamente os textos relevantes de Jack Balkin, John Bowers e Jonathan Zittrain, os padrões internacionais para a regulação de plataformas digitais, como os Princípios de Manila (2015), de Santa Clara (2018), e o inteiro teor do PL 2.630/2020.

Com isso em mente, o Capítulo 1 buscará inicialmente delinear o conceito de liberdade de expressão, seu enquadramento no ordenamento jurídico brasileiro e no direito internacional. Em seguida, serão expostos os novos contornos regulatórios da liberdade de expressão com a chegada da internet. Logo após, será introduzido o modelo triangular de regulação da liberdade de expressão proposto por Jack Balkin (2018), além dos desafios e soluções concebidas pelo autor. Ato contínuo, apresentar-se-á o esquema de eras regulatórias da governança de conteúdo online, desenvolvido por Bowers e Zittrain (2020). Finalmente, a partir dos consensos identificados entre os autores de referência, e as recomendações elencadas em padrões internacionais, serão dispostos os elementos constitutivos da abordagem regulatória procedimental.

O Capítulo 2, por sua vez, buscará identificar, por meio de uma pesquisa documental, quais mecanismos e inovações regulatórias típicas da abordagem processual estão previstos no texto do Projeto de Lei nº 2.630 aprovado pelo Senado Federal em junho de 2020 (sem considerar versões futuras). Também será feita uma contextualização geral da proposta, sua estrutura, e processo legislativo.

Por fim, o Capítulo 3 analisará os dispositivos do PL 2.630/20 previamente discriminados na pesquisa documental, à luz da *perspectiva procedimental* de regulação das plataformas digitais. Serão objeto de consideração detida os instrumentos regulatórios de caráter procedimental previstos pelo legislador pátrio, sua técnica legislativa, aspectos positivos, negativos e possíveis efeitos, caso a proposta em apreço seja aprovada.

⁷ “[...] por autorregulação deve-se entender a existência de regras comportamentais desenvolvidas com alguma participação do próprio destinatário ou exclusivamente desenvolvidas por ele” (ARANHA, 2019).

1.1 NOTAS METODOLÓGICAS

Todas as referências feitas ao texto do PL 2.630/2020 neste trabalho dizem respeito à versão aprovada pelo Senado Federal em 30 de junho de 2020. Isso porque a presente pesquisa se deu nos meses iniciais de 2021, momento em que este era o único documento público em tramitação no âmbito dessa iniciativa legiferante. Portanto, não foram consideradas outras versões, relatórios, substitutivos ou alterações posteriores feitas à referida proposta legislativa pelo Congresso Nacional.

2 UMA PERSPECTIVA PROCEDIMENTAL PARA REGULAR A LIBERDADE DE EXPRESSÃO NA INTERNET

O objetivo deste capítulo é descrever os fundamentos jurídicos e teóricos para uma perspectiva procedimental de regulação do direito à liberdade de expressão na era digital, tendo como principal eixo o modelo conceitual desenvolvido pelo professor Jack M. Balkin (*Yale Law School*). Em diálogo com essa perspectiva, também é apresentado o esquema de eras regulatórias da governança privada de conteúdo proposto por John Bowers e Jonathan Zittrain (*Harvard Law School*). Conquanto os autores escolhidos para compor o referencial teórico estejam inseridos no contexto norte-americano, buscou-se correlacionar suas ideias com a realidade brasileira e com os padrões internacionais de direitos humanos.

Este capítulo foi dividido em cinco subcapítulos, seguindo uma continuidade lógica e didática sobre o tema. A primeira parte introduz o conceito de liberdade de expressão, e sua aceção no direito brasileiro e internacional. Em seguida, delineia-se os novos contornos regulatórios da liberdade de expressão com a chegada da internet. Logo após, é apresentado o modelo pluralista de regulação da liberdade de expressão proposto por Balkin, seus problemas e as soluções regulatórias propostas pelo autor. A quarta seção explora as três eras da regulação de plataformas digitais descritas por Bowers e Zittrain. E por último, é feita uma breve recapitulação dos elementos que caracterizam a abordagem regulatória direcionada ao processo, tomando como base os consensos e as sugestões trazidas pelos autores de referência.

2.1 A LIBERDADE DE EXPRESSÃO COMO DIREITO FUNDAMENTAL

A liberdade de expressão, na aceção deste trabalho, é compreendida como o direito fundamental subjetivo, assegurado a qualquer indivíduo, “de manifestar livremente os próprios pensamentos, ideias, opiniões, crenças, juízos de valor, por meio da palavra oral e escrita, da imagem, ou de qualquer outro meio de difusão” (FARIAS, 2001, p, 45).

Em sentido amplo, a liberdade de expressão abrange todas as liberdades de comunicação, alcançando as mais variadas formas de expressão humana. Isso inclui não apenas sua dimensão subjetiva, entendida como o direito à livre manifestação do pensamento e opinião⁸, mas também as liberdades de criação, de imprensa, e o direito à informação (TÔRRES, 2013, p. 62). A liberdade de comunicação também designa uma função social, a saber, a liberdade de “comunicar ou receber informações verdadeiras, sem impedimentos nem discriminações” (FARIAS, 2001, p. 45).

Trata-se, portanto, de uma salvaguarda ampla que não se limita à proteção da dignidade humana em sua autonomia individual, mas que alcança a própria coletividade. Isso porque constitui um valor indispensável ao funcionamento da estrutura democrática, a qual pressupõe o pluralismo de ideias, a livre formação de opiniões, e a participação ativa de todos no debate público (TÔRRES, 2013, p. 61).

Dessa constatação decorrem duas importantes tarefas às democracias constitucionais contemporâneas. Por um lado, há a necessidade de assegurar a livre circulação de ideias, pensamentos, fatos, e opiniões na vida social. Por outro, requer-se que os abusos cometidos no exercício das liberdades de expressão sejam enfrentados a fim de proteger os cidadãos (FARIAS, 2001, p. 290).

Tais abusos, como bem se sabe, não se restringem aos órgãos de comunicação em massa, tipicamente identificados com a imprensa (ex.: televisão, rádio, jornal etc.). Conquanto os meios tradicionais de difusão de informação protagonizaram o exercício das liberdades comunicativas ao longo dos séculos XIX e XX, esta dinâmica se modificou a partir da década de 90 com a popularização da internet e a chegada das mídias digitais (BALKIN, 2018, p. 2011).

Ao longo das últimas décadas, a internet passou a ser percebida como um “bem público global”, sendo também designada como a nova praça pública⁹ na cultura

⁸ A dimensão subjetiva da liberdade de expressão, própria da autonomia pessoal, e que designa a faculdade de “*externar ideias, opiniões, juízos de valor e [...] qualquer manifestação do pensamento humano*” é entendida como a liberdade de expressão *em sentido estrito*. Este trabalho monográfico concentra-se, sobremaneira, nesta acepção específica (BARROSO, 2004, p. 18)

⁹ Conforme detalhado por Al Sur (2021, p. 7), a doutrina da praça pública se origina no precedente *Hague vs. CIO* (1939) da Suprema Corte estadunidense. Nessa decisão, a corte daquele país considerou que a proibição de encontros “com teor político em espaços públicos” deveria ser considerada inconstitucional por violarem a Primeira Emenda. Essa decisão reverteu o precedente que “concedia poder aos funcionários públicos sobre os espaços públicos como o de proprietários sobre a propriedade privada e impõe a obrigação de facilitar e proteger a liberdade de expressão nos espaços públicos. *Hague v. CIO, 307 US at 515-16*”. Com o passar do tempo, essa doutrina foi atualizada ao

jurídica norte-americana (STERN *et al*, 1999; KAUL *et al*, 2003; CANAZZA, 2018; AL SUR, 2021, p. 7). À medida que as redes sociais passaram a desempenhar um papel central nos processos comunicacionais a nível global, modificou-se também o modo como se dava o exercício da liberdade de expressão (BALKIN, 2018, p. 2011).

Neste novo cenário, as plataformas digitais não são meras intermediárias neutras do processo comunicativo. Antes, porém, intervêm diretamente sobre este, exercendo o papel de verdadeiros governantes privados desse espaço virtual. Essas empresas criam e aplicam seus próprios termos de serviço e padrões da comunidade, definindo o que é ou não permitido em seus domínios (KLONICK, 2018, p. 1603; FRAZÃO, 2018, p. 635).

Por tomarem decisões constantes e ubíquas sobre o conteúdo gerado por seus usuários, as plataformas atuam também como *curadoras* ou *moderadoras de conteúdo*. Para isso, empregam uma combinação de algoritmos e colaboradores humanos, inseridos em uma hierarquia corporativa cujo funcionamento é, em geral, pouco transparente. Todos esses elementos compõem o quadro da governança privada da liberdade de expressão, que marca esse novo contexto (BALKIN, 2018, p. 2041).

Sob essa nova realidade, que será aprofundada mais adiante, a concretização do direito fundamental à liberdade de expressão demanda uma abordagem regulatória distinta, capaz de endereçar os problemas e relações disfuncionais que a caracterizam. Para isso, é preciso reinterpretar esse bem jurídico tão caro à democracia, aplicando-o à essa nova lógica comunicativa, de modo a garantir sua justa e efetiva proteção na esfera digital.

2.1.1 A liberdade de expressão no direito brasileiro

A Constituição Federal de 1988 concede ampla proteção à liberdade de expressão em suas várias expressões. Além de inseri-la no rol de direitos e garantias fundamentais (art. 5º, incisos IV, V, X, XIII e XIV), o constituinte dedicou um capítulo exclusivo para tratar da “Comunicação Social” (arts. 220 a 224). No *caput* do art. 220,

por diversas outras decisões que a aplicaram ao ambiente virtual, estendendo-a até mesmo às redes sociais na decisão paradigmática de “*Packingham vs. North Carolina*” (2017, p. 1735-36; STARZAK, 2017; LANE, 2019).

a Carta Política assegura que “a manifestação do pensamento, a criação, a expressão, e a informação, sob qualquer forma, processo ou veículo não sofrerão qualquer restrição, observado o disposto no art. 5º, IV, V, X, XIII e XIV”. Acrescenta, nos §§ 1º e 2º do mesmo artigo, que “nenhuma lei conterà dispositivo que possa constituir embaraço à plena liberdade de informação jornalística em qualquer veículo de comunicação social”, e veda, de forma peremptória, “toda e qualquer censura de natureza política, ideológica e artística” (BRASIL, 1988).

Ainda, o texto constitucional relaciona a liberdade de expressão com outros direitos conexos, quais sejam o direito de resposta e de réplica política (art. 5º, inciso V), o direito de informar e ser informado (art. 5º, inciso X), a liberdade de reunião (art. 5º, inciso XVI), a liberdade religiosa (art. 5º, inciso VI e VIII), entre outros (TÔRRES, 2013, p. 63).

A ênfase do constituinte em proteger a liberdade de expressão tem uma razão histórica. Naquele momento, o Brasil encerrava uma ditadura militar que praticou abertamente a censura contra a imprensa, artistas e intelectuais. Episódios nefastos como a invasão, depredação e interdição de jornais críticos ao regime¹⁰, foram sucedidos pela prática sistemática da censura prévia¹¹ na redação dos jornais, sujeitando cada publicação ao crivo dos censores (OLIVIERI, 2014; BARROSO, 2021, p. 3). Por quase vinte anos, conteúdos considerados perigosos ou ofensivos aos interesses do regime poderiam ser objeto de reprimenda e censura governamental. A nova constituição deveria lançar as bases necessárias para que o país de se desenvolvesse de forma livre e democrática (SARMENTO, 2006, p. 46).

Com efeito, o direito fundamental à liberdade de expressão carrega consigo uma pretensão de caráter defensivo, voltada a impedir que o Estado pratique a censura. Ou seja, o ente estatal deve abster-se de interferir na esfera de liberdade de

¹⁰ Durante a primeira fase da censura no regime militar (1964-1968), jornais de esquerda, periódicos próximos a João Goulart e noticiários críticos aos militares foram alvo de depredações, invasões e interdições. Dentre os veículos de imprensa atacados estão a “Folha da Semana”, o “Última Hora” e o “Correio da Manhã”. Neste período também foram estabelecidas limitações à liberdade de imprensa com a Lei nº 5.250, de 9 de fevereiro de 1967, denominada Lei de Imprensa. (OLIVIERI, 2014).

¹¹ A segunda e mais implacável fase de censura na ditadura militar teve início com o AI-5, agravou-se com o Decreto-Lei nº 898 de 29 de setembro de 1969, conhecido como Lei de Segurança Nacional (LSN), e aperfeiçoou-se com o Decreto-Lei nº 1.077 de 1970 que instituiu a censura prévia. Este último previa a presença permanente de censores nas redações do país a fim de autorizar ou vedar certos conteúdos antes de sua veiculação (OLIVIERI, 2014).

manifestação dos cidadãos, conforme disposto no art. 220 do texto constitucional (MENDES & BRANCO, 2015, p. 264-265). Isso não impede, contudo, que o poder público regulamente tal direito visando *ampliar* sua efetivação (TÔRRES, 2013, p. 64).

Destarte, não é papel do Estado definir quais opiniões são mais ou menos valorosas, aceitáveis, ou merecedoras de crédito. Esta faculdade pertence à própria sociedade, que, em meio a um debate amplo, livre e público de ideias poderá, por si só, abraçar ou rejeitar as visões que julgar merecedoras, dentro dos limites legais¹². Tal compreensão está na própria raiz da noção de uma democracia livre e plural (KARPEN, 1988, pg. 93).

Corroborando esse entendimento, o Supremo Tribunal Federal no âmbito da ADI 4451, declarou a inconstitucionalidade de alguns trechos da Lei Federal nº 9.504/97, também conhecida como Lei das Eleições. Os dispositivos impugnados proibiam a veiculação de peças com tom humorístico que criticassem candidatos, partidos ou coligações (BRASIL, 2019). No acórdão, o relator argumenta que:

“A Democracia não existirá e a livre participação política não florescerá onde a liberdade de expressão for ceifada, pois esta constitui condição essencial ao pluralismo de ideias, que por sua vez é um valor estruturante para o salutar funcionamento do sistema democrático. A livre discussão, a ampla participação política e o princípio democrático estão interligados com a liberdade de expressão, tendo por objeto não somente a proteção de pensamentos e ideias, mas também opiniões, crenças, realização de juízo de valor e críticas a agentes públicos, no sentido de garantir a real participação dos cidadãos na vida coletiva [...] Tanto a liberdade de expressão quanto a participação política em uma Democracia representativa somente se fortalecem em um ambiente de total visibilidade e possibilidade de exposição crítica das mais variadas opiniões sobre os governantes.” (ADI nº 4.451, Rel. Min. Alexandre de Moraes, julgamento em 21/06/2018, Plenário, DJE de 06/03/2019).

¹² Essa ressalva ecoa o “paradoxo da tolerância”, enunciado pelo filósofo Karl Popper, segundo o qual, em nome da tolerância, dever-se-ia reservar o “direito de não tolerar o intolerante”. O autor austro-britânico defende uma ampla acepção do termo “liberdade de expressão”, e sua proteção aos variados discursos que povoam o debate público. No entanto, essa tolerância não deveria ser ilimitada. Popper afirma que as “filosofias intolerantes”, devem, em primeiro lugar, serem enfrentadas com argumentos, no bojo do próprio debate público. Há, contudo, situações nas quais seria possível recorrer a um direito de “autopreservação”, inerente ao próprio sistema democrático (POPPER, 1945, p. 118-119).

Não se trata, contudo, de um direito absoluto. Na ocorrência de um conflito com outros direitos fundamentais ou mesmo com valores caros à Constituição, é possível que um deles sofra limitações no caso concreto (CANOTILHO, 2003, p. 1276).

Tais restrições podem ocorrer por força da própria Constituição, ou caso esta permita, por uma lei infraconstitucional. Não é por acaso que a carta magna contrapõe à liberdade de expressão, outros direitos fundamentais como o direito à “inviolabilidade da intimidade, vida privada, honra e imagem das pessoas” (art. 5º, X) e o direito à “indenização por dano moral ou à imagem” (art. 5º, V). Há, ainda, outros bens e valores constitucionais que podem colidir com a liberdade de expressão, como a vida, a integridade e o devido processo legal etc. (SARMENTO, 2006, p. 46).

Soma-se a esse conjunto de garantias com guarida constitucional, o direito fundamental à proteção de dados pessoais, reconhecido pelo plenário do Supremo Tribunal Federal em julgamento cautelar da Medida Provisória n. 954/2020 no âmbito da Ação Direta de Inconstitucionalidade nº 6390 (BRASIL, 2020). A decisão histórica consagrou a proteção de dados como um direito fundamental autônomo, distinto da proteção à intimidade e à privacidade (SCHERTEL, 2020). Trata-se, portanto, da admissão de um novo direito com status constitucional que traz repercussões diretas nas questões concernentes à liberdade de expressão no âmbito digital.

Não obstante o reconhecimento deste novel direito pelo pleno do Supremo Tribunal Federal, o Congresso Nacional recentemente aprovou a Proposta de Emenda à Constituição (PEC) nº 17/2019, que inclui o direito à proteção de dados no rol de direitos e garantias fundamentais do art. 5º da Constituição Federal. O texto final também fixou a competência privativa da União para legislar sobre proteção e tratamento de dados pessoais (BRASIL, 2021).

Conquanto a Lei Maior estabeleça uma ampla guarida ao exercício da liberdade de expressão, essa proteção não alcança certas práticas como a incitação à violência e outros discursos de ódio¹³. A liberdade de expressão, portanto, não configura uma

¹³ Embora não haja consenso doutrinário quanto à definição de “discurso de ódio”, é possível afirmar que seu significado abrange “manifestações de ódio, desprezo ou intolerância contra determinados grupos, motivados por preconceitos ligados à etnia, religião, gênero, deficiência física ou mental e orientação sexual, dentre outros” (SARMENTO, 2006, p. 46). Há ainda uma grande quantidade de condutas que poderiam ser consideradas discursos de ódio, como calúnias e insultos. Estes, contudo, trazem consigo a dimensão subjetiva da honra, que pode dar azo a divergências sobre sua concretização ou não. Também estão inclusos nessa definição os discursos violentos, que promovem,

concessão de imunidade, nem tampouco um salvo-conduto em relação àquilo que se diz. Antes, porém, a liberdade é uma condição para a responsabilização, e não a isenção dela, razão pela qual a CFRB/88 veda o anonimato. O “emissor” da mensagem no processo comunicativo deve assumir as consequências cíveis e penais daquilo que expressou (REALE JÚNIOR, 2010, p. 382; MENDES *et al*, 2015, p. 265; CGI, 2020, p. 13).

O próprio Supremo Tribunal Federal tem confirmado esse entendimento em julgamentos¹⁴ que enquadraram manifestações de índole antissemita no crime de incitação ao racismo. A Corte também equiparou a homofobia ao racismo, em caso de “ameaças, agressões e discriminações” contra indivíduos da comunidade LGBTQIA+ (BARROSO, 2020, p. 9).

Quando associadas à internet e às mídias sociais¹⁵, certas práticas como a promoção de discursos de ódio, campanhas de desinformação¹⁶ e ataques às instituições democráticas ganham uma escala, alcance e potencial de dano nunca

incentivam ou causem o dano físico contra terceiros ou até mesmo os que incitem a violência em massa. Esta violência, contudo, não se restringe à violência armada ou física, mas também pode caracterizar uma violência psicológica (RUEDIGER *et al*, 2021, p. 12-13). É importante destacar, todavia, que palavras duras, desagradáveis, ou mesmo reprováveis não estão necessariamente fora da salvaguarda da liberdade de expressão, de modo que os limites do discurso devem ser analisados no caso concreto (MENDES & BRANCO, 2015, p. 273-274).

¹⁴ Em 2003, o Supremo Tribunal Federal proferiu um julgamento paradigmático sobre os discursos de ódio no caso *Ellwanger* (BRASIL, 2003). Discutiu-se na ocasião a publicação de material discriminatório contra o povo judeu, voltado a avançar teses revisionistas que negavam a ocorrência do holocausto durante a Segunda Guerra Mundial, e relativizavam o horror nazista (VIOLANTE, 2010, p. 15). Na decisão, os ministros consideraram que a liberdade de expressão não resguardava a veiculação de conteúdo com índole antissemita, sujeitas à persecução penal por crime de racismo (SARMENTO, 2006, p. 3). O tribunal adotou a mesma linha de entendimento em relação a ameaças, agressões e discriminações contra pessoas da comunidade LGBTQIA+, equiparando a homofobia ao racismo (BARROSO, 2020, p. 9).

¹⁵ Na era digital, práticas como a desinformação, a propagação de discursos de ódio e a incitação à violência ganharam um alcance e dimensão sem precedentes na história. Isso porque a internet possibilita a criação e a disseminação de conteúdo com enorme rapidez, abrangência e de forma, muitas vezes, anônima. Mais do que isso, as ferramentas tecnológicas à disposição na atualidade permitem o direcionamento de anúncios de forma individualizada, ajustados aos interesses e o perfil psicológico de cada usuário. Essa personalização é capaz de potencializar o impacto das mensagens veiculadas. A combinação dessas novas condições criou uma “tempestade perfeita” de informações falsas, intolerância, discriminação e violência no ambiente virtual (LAPIN, 2020, p. 3).

¹⁶ O conceito de “desinformação” abrange “todas as formas de informação falsa, imprecisa ou enganosa, desenhadas, apresentadas e promovidas *intencionalmente para causar dano público ou gerar lucro*” (EC, 2018, p. 10, grifo do autor). É, contudo, necessário desambiguar o termo desinformação (disinformation) dos verbetes “misinformation”, que designa “informações falsas que são disseminadas, *independentemente da intenção de induzir em erro*”, e a expressão “fake news”, que corresponde a “informações propositalmente elaboradas, sensacionalistas, emocionalmente carregadas, enganosas ou totalmente fabricadas que imitam a forma de notícias convencionais” (ZIMDARS *et al*, 2020; DICTIONARY.COM, 2021)

antes visto. Essas novas questões têm desafiado empresas de internet, governos, políticos e tribunais na busca por soluções (BARROSO, 2020, p. 10-11).

2.1.1.1 A responsabilização civil por conteúdo gerado por terceiros

Na era digital, não obstante, a responsabilização legal ganha contornos peculiares. Isso porque, nesse novo contexto, os provedores de aplicação de internet (como Facebook, Twitter, Youtube etc.) moldam, estruturam e mediam o modo pelo qual as pessoas interagem, emitem opiniões e se relacionam umas com as outras, além de interferir diretamente na forma como os conteúdos são produzidos, distribuídos e consumidos na internet (VERONESE *et al*, 2018, p. 43). Dada a sua centralidade no processo comunicativo, discute-se, desse modo, se esses intermediários devem ser solidariamente responsabilizados ou não por eventuais danos causados por conteúdos publicado por terceiros em seus espaços virtuais.

O art. 19¹⁷ do Marco Civil da Internet (MCI) estabeleceu, como regra geral, um regime de exclusão da responsabilidade dos provedores. Há, contudo, duas exceções à imunidade de responsabilização previstas pelo MCI. A primeira diz respeito a conteúdos que violem direitos autorais. Nesse caso, os provedores de aplicações somente serão passíveis de responsabilização quando, em sendo notificados de decisão judicial específica determinando a sua remoção, não tomarem providências para tornar indisponível o conteúdo apontado como ilícito pela autoridade judicial (BRASIL, 2014; OLIVA, 2019). Trata-se, portanto, de uma dinâmica de responsabilidade civil a partir de descumprimento de ordem judicial (“*judicial notice and take down*”) (SOUZA *et al*, 2016; FONSECA, 2019, p. 65 *apud* VERONESE *et al*, 2017).

Na prática, contudo, as empresas adotam um sistema similar ao instituído nos EUA assentado em sua Lei de Direitos Autorais na Era Digital (*Digital Millennium*

¹⁷ Artigo 19 do Marco Civil da Internet: “Com o intuito de assegurar a liberdade de expressão e impedir a censura, o provedor de aplicações de internet somente poderá ser responsabilizado civilmente por danos decorrentes de conteúdo gerado por terceiros se, após ordem judicial específica, não tomar as providências para, no âmbito e nos limites técnicos do seu serviço e dentro do prazo assinalado, tornar indisponível o conteúdo apontado como infringente, ressalvadas as disposições legais em contrário.” (BRASIL 2014).

Copyright Act, ou, “DMCA”), que estabelece um modelo de “notificação e retirada” (*notice and take down*). Nessa dinâmica, uma simples notificação privada (“extrajudicial”) do proprietário afetado é suficiente para que o provedor de serviços intermediários remova ou bloqueie o conteúdo supostamente infrator (AL SUR, 2021, p. 12, 19). Isso acontece muito em razão dessas companhias serem, em sua grande maioria, sediadas naquele país, sujeitando-se à sua cultura jurídica e influência.

A segunda exceção incide sob a hipótese de divulgação não consensual de imagens íntimas. Neste caso, bastará uma notificação do próprio usuário para que a plataforma remova o conteúdo. Caso contrário, o provedor de aplicação da internet poderá ser responsabilizado subsidiariamente por violação da intimidade (AL SUR, 2021, p. 12; SANTOS, 2020).

Há, contudo, uma diferença importante entre a legislação brasileira e a norte-americana no que diz respeito à responsabilização por conteúdo de terceiros e à moderação de conteúdo. Nos EUA, a Seção 230¹⁸ da Lei de Decência das Comunicações (*Communications Decency Act*, ou “CDA”), aprovada em 1996, estabelece duas provisões importantes a esse respeito.

A primeira delas, conhecida como cláusula do “porto seguro” (*safe harbour*), prevê como regra geral a não responsabilização dos provedores por conteúdo gerado por outrem. Tal qual o modelo brasileiro, esse dispositivo protege esses atores em relação aos discursos propagados por usuários em suas redes. Na época de sua criação, traçou-se um paralelo entre o papel das plataformas digitais e o das empresas de telefonia. Isso porque estas forneciam o meio para a comunicação, mas quem se responsabilizava pelas mensagens transmitidas eram os usuários (GILLESPIE, 2018).

Em seguida, a Seção 230 garante imunidade aos intermediários caso estes, de boa-fé, removam conteúdo de terceiros de suas plataformas ainda que este seja protegido constitucionalmente. É a chamada “cláusula do bom samaritano”¹⁹ (AL

¹⁸ Em tradução livre, a Seção 230 prevê que “nenhum provedor ou usuário de um serviço interativo de computador deve ser tratado como o editor ou locutor de qualquer informação fornecida por outro provedor de conteúdo informativo”, sendo dessa forma imune a qualquer responsabilização decorrente do conteúdo produzido por terceiros (EUA, 1996).

¹⁹ CDA Section 230 “[...] (c) *Protection for “Good Samaritan” blocking and screening of offensive material* (1) *Treatment of publisher or speaker* No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider. (2) *Civil liability* No provider or user of an interactive computer service shall be held liable on account of— (A) any action voluntarily taken in good faith to restrict access to or availability of material

SUR, 2021, p. 16; SANTOS, 2020). Tal dispositivo permite expressamente a moderação de conteúdo na internet, incentivando os provedores a assumir uma postura proativa de combate à circulação de material danoso ou obsceno em seus espaços. Essa ampla proteção legal contra a responsabilização impactou significativamente o desenvolvimento da internet, e as práticas de moderação de conteúdo adotadas pelas grandes plataformas (CITRON *et al*, 2017, p. 408; KLONICK, 2018, p. 1602; IRIS, 2020).

A principal crítica ao modelo norte-americano que fornece um “porto seguro” contra a responsabilização por conteúdo de terceiros, além de dar aos provedores o direito de moderar conteúdos, é que ele não protege suficientemente os usuários contra o poder quase absoluto de moderação das plataformas digitais. Tais empresas tomam decisões constantes e significativas sobre o conteúdo gerado por seus usuários, exercendo enorme controle sobre o que estes veem, e sobre as informações que circulam em suas comunidades digitais. Conquanto as provisões legais de imunidade tenham sido criadas para proteger o papel dos intermediadores de informação da internet, as plataformas digitais não são, hoje, mediadores neutros do processo comunicativo (FRAZÃO *et al*, 2021).

No Brasil, porém, o Marco Civil da Internet não proíbe que os provedores sejam responsabilizados por atos próprios de moderação de conteúdo. Isto configura, por si só, uma garantia ao usuário, que pode acionar o Poder Judiciário caso se sinta lesado por uma ação de moderação da plataforma. É o que tem acontecido, por exemplo, no âmbito do Tribunal de Justiça de São Paulo (TJ-SP), que já proferiu inúmeras decisões sobre as práticas de moderação de conteúdo, muitas vezes em desfavor das próprias plataformas. Portanto, o art. 19 é compatível²⁰ tanto com a remoção de conteúdo via ordem judicial quanto com a moderação pelas plataformas de acordo com os seus

that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected; or (B) any action taken to enable or make available to information content providers or others the technical means to restrict access to material described in paragraph (1)” (EUA, 1996).

²⁰ Carlos Affonso de Souza (2021) defende que o “Judiciário deve sempre dar a última palavra sobre o que é postado na internet com base na CFRB/88 e nas leis nacionais, mas a primeira palavra sobre conteúdos danosos pode/deve ser das plataformas por meio de seus termos de uso. Se esses termos de uso forem abusivos, ou sua aplicação for errônea, abusiva ou ilícita, o Poder Judiciário deve ser acionado [...] Desse modo, no modelo que temos hoje, os termos de uso constituem uma primeira resposta a conteúdo problemático, e o acesso à justiça é o caminho para que buscar uma última palavra [...] Portanto, o Judiciário deve funcionar como uma rede de proteção, a ser acionada em casos difíceis, em que há discordância sobre a decisão de moderação”.

termos de uso, os quais podem ser questionados judicialmente e não devem violar as normas da CFRB/88. (SOUZA & TEFFÉ, 2021).

De fato, o Marco Civil da Internet não trata diretamente da moderação de conteúdo. Isso porque a norma foi aprovada em 2014, época na qual esse tema ainda não ocupava um papel de centralidade no debate público. A título de contextualização, os padrões da comunidade do Facebook foram publicados apenas em 2018, após o vazamento na imprensa de documentos internos da empresa com instruções sobre moderação de conteúdo (ARHEGAS *et al*, 2021, p. 6). Isso explica a falta de provisões específicas sobre o tema na referida lei.

2.1.2 A liberdade de expressão no direito internacional

Após a Segunda Guerra Mundial, diversos tratados internacionais sobre direitos humanos foram celebrados, prevendo regras e garantias à liberdade de expressão. A Declaração Universal dos Direitos Humanos (DUDH), documento fundacional para o direito internacional humanitário, garante²¹ a todos o direito de, sem interferências, expressar opiniões e de buscar, receber e transmitir informações e ideias de toda espécie, independentemente de fronteiras, ou meios (SARMENTO, 2006, p. 25). O mesmo estatuto, por sua vez, assinala em seu art. 29 que o exercício dos direitos e liberdades fundamentais ali previstos estão condicionados “aos deveres para com a comunidade, assim como aos direitos e liberdades dos outros”, conforme elucida Lindgren Alves (1997 *apud* LAFER, 2004, p. 75)

De modo semelhante, o Pacto Internacional sobre Direitos Cívicos e Políticos (PIDP), do qual o Brasil é signatário, ecoa a previsão do DUDH, e garante, em seu décimo nono artigo²², amplo direito à liberdade de expressão (ONU, 2018, p. 4). O

²¹ Art. 19 da Declaração Universal de Direitos Humanos “Todo o indivíduo tem direito à liberdade de opinião e de expressão, o que implica o direito de não ser inquietado pelas suas opiniões e o de procurar, receber e difundir, sem consideração de fronteiras, informações e ideias por qualquer meio de expressão.” (ONU, 1948)

²² “Artigo 19 - 1. Ninguém poderá ser molestado por suas opiniões. 2. Toda pessoa terá direito à liberdade de expressão; esse direito incluirá a liberdade de procurar, receber e difundir informações e ideias de qualquer natureza, independentemente de considerações de fronteiras, verbalmente ou por escrito, em forma impressa ou artística, ou por qualquer outro meio de sua escolha. 3. O exercício do direito previsto no parágrafo 2 do presente artigo implicará deveres e responsabilidades especiais. Consequentemente, poderá estar sujeito a certas restrições, que devem, entretanto, ser expressamente previstas em lei e que se façam necessárias para: a) assegurar o respeito dos direitos

mesmo dispositivo, mais adiante, prevê a possibilidade de que sejam estabelecidas restrições à liberdade de expressão. Para tanto, exige o cumprimento de três condições: (i) legalidade, (ii) necessidade e proporcionalidade, e (iii) legitimidade.

O quesito da legalidade requer que quaisquer restrições à liberdade de expressão sejam “expressamente previstas em lei” (BRASIL, 1992). Isso inclui não apenas a existência prévia de normas que disciplinem tal medida, mas também a adoção dos processos legais adequados, capazes de limitar a discricionariedade estatal. Em geral, a garantia da legalidade pressupõe também o controle judicial exercido por autoridades independentes (ONU, 2018, p. 4).

Em segundo lugar, a imposição de limites à liberdade de expressão deve atender aos requisitos de necessidade e proporcionalidade. Isso significa que para a aplicação de uma medida restritiva, o Estado deve demonstrar que aquela providência é adequada, proporcional e absolutamente necessária para garantir a proteção de outro bem jurídico legítimo. Ou seja, não basta uma mera decisão, é preciso que as razões que motivaram aquele juízo sejam explicitadas e demonstrem a pertinência e a imprescindibilidade daquela medida (ONU, 2018, p. 4).

Por último, é necessário que as restrições ao referido direito sejam legítimas, e, portanto, visem proteger unicamente os interesses jurídicos elencados no Artigo 19. Incluem-se aí situações em que seja preciso assegurar o respeito aos direitos e à reputação das demais pessoas²³, ou ainda proteger outros valores legítimos como a segurança nacional, a ordem, a saúde ou a moral públicas (BRASIL, 1992).

É possível ainda mencionar outros documentos firmados pela comunidade internacional que resguardam o direito à liberdade de expressão. Dentre eles estão a Convenção Interamericana de Direitos Humanos (art. 13)²⁴, também conhecida como

e da reputação das demais pessoas; b) proteger a segurança nacional, a ordem, a saúde ou a moral públicas.” (BRASIL, 1992)

²³ Por “direitos [...] das demais pessoas” compreende-se não apenas os direitos reconhecidos pelo próprio PIDP, mas também aqueles resguardados pelo direito internacional humanitário (ONU, 2011, p. 7). Aí encontram-se, por exemplo, os direitos à privacidade, à vida, à associação e participação no debate público, ao devido processo, entre outros (ONU, 2018, p. 4).

²⁴ Artigo 13. Liberdade de pensamento e de expressão 1. Toda pessoa tem direito à liberdade de pensamento e de expressão. Esse direito compreende a liberdade de buscar, receber e difundir informações e ideias de toda natureza, sem consideração de fronteiras, verbalmente ou por escrito, ou em forma impressa ou artística, ou por qualquer outro processo de sua escolha. 2. O exercício do direito previsto no inciso precedente não pode estar sujeito a censura prévia, mas a responsabilidades ulteriores, que devem ser expressamente fixadas pela lei e ser necessárias para assegurar: (a). o respeito aos direitos ou à reputação das demais pessoas; ou (b). a proteção da segurança nacional, da

Pacto de San José da Costa Rica, a Convenção Europeia de Direitos Humanos (art. 10), e a Carta Africana de Direitos Humanos (art. 9º).

2.2 OS NOVOS CONTORNOS REGULATÓRIOS DA LIBERDADE DE EXPRESSÃO NA ERA DIGITAL

Do início do século XIX ao fim do século XX, prevaleceu o modelo dualista de regulação da liberdade de expressão. Dualista, pois, se constituía em uma relação com dois polos. Em um dos lados estava o poder público, na figura do Estado, que buscava ordenar e regular o discurso público em maior ou menor intensidade. O outro lado era ocupado pelos “emissores” (*speakers*), ou “locutores”, que são os sujeitos ativos no processo comunicativo. Os emissores podem ser tanto os veículos da imprensa tradicional, quanto indivíduos no exercício de suas liberdades comunicativas, como artistas, intelectuais etc. (BALKIN, 2018, p. 2015).

Esta velha dinâmica de regulação da liberdade de expressão caracteriza o que Balkin denomina como *abordagem tradicional de regulação da liberdade de expressão* (“*old-school speech regulation*”) (BALKIN, 2018, p. 2015).

Nessa abordagem antiga, ou tradicional, o Estado recorre a instrumentos como multas, penalidades, aprisionamento, dentre outras formas de punição e retribuição com o fim de exercer controle e efetivamente regular o discurso de indivíduos ou da mídia (BALKIN, 2018, p. 2015-2016).

Tal abordagem regulatória ainda é praticada em todo o mundo, inclusive no Brasil. É o que acontece, por exemplo, quando uma revista veicula crítica jornalística com intuito de difamar, injuriar ou caluniar algum indivíduo. Nesse caso, há uma violação aos deveres de cuidado, pertinência pública e veracidade intrínsecos à

ordem pública, ou da saúde ou da moral públicas. 3. Não se pode restringir o direito de expressão por vias ou meios indiretos, tais como o abuso de controles oficiais ou particulares de papel de imprensa, de frequências radioelétricas ou de equipamentos e aparelhos usados na difusão de informação, nem por quaisquer outros meios destinados a obstar a comunicação e a circulação de ideias e opiniões. 4. A lei pode submeter os espetáculos públicos a censura prévia, com o objetivo exclusivo de regular o acesso a eles, para proteção moral da infância e da adolescência, sem prejuízo do disposto no inciso 2. 5. A lei deve proibir toda propaganda a favor da guerra, bem como toda apologia ao ódio nacional, racial ou religioso que constitua incitação à discriminação, à hostilidade, ao crime ou à violência.

atividade da imprensa. Em função disso, podem ser responsabilizados por dano em decorrência de ilícito. A própria CFRB/88 garante o direito à resposta e reparação nesses casos (MARINELLI, 2020).

Ocorre que, ao longo das últimas décadas, os fluxos de comunicação deslocaram-se cada vez mais para o domínio digital. Nesse novo contexto, há um novo sujeito que desempenha papéis cruciais. Além de moldar, estruturar e mediar as relações comunicativas, as empresas de internet passaram a controlar os meios pelos quais as pessoas interagem, emitem opiniões e se relacionam umas com as outras. Além disso, interferem diretamente no modo como conteúdos são produzidos, distribuídos e consumidos (VERONESE *et al*, 2018, p. 43).

Nesse novo cenário, próprio ao século XXI, foram desenvolvidas novas técnicas de regulação com o objetivo de controlar o discurso em um ambiente dominado por empresas de internet. Esse novo ferramental regulatório caracteriza o que Balkin chama de *nova abordagem de regulação da liberdade de expressão (new-school speech regulation)* (BALKIN, 2018, p. 2015-2016).

Enquanto a antiga abordagem (*old-school*) é voltada aos emissores (*speakers*), ou seja, aos criadores de conteúdo, a nova abordagem (*new-school*) direciona-se às empresas de internet. Com a nova estratégia, o poder público busca “ameaçar, coagir ou cooptar a infraestrutura da internet com o objetivo de vigiar, policiar e controlar os emissores” (BALKIN, 2018, p. 2016).

Dentre as características mais distintivas da nova abordagem está a tentativa do Estado de “terceirizar” ou “delegar” algumas de suas funções públicas. Ao invés de regular diretamente os emissores, como fazia antes, agora o Estado os regula indiretamente²⁵, visto que esses novos espaços “públicos” de comunicação são

²⁵ A cooperação e a cooptação das plataformas pelo Estado decorrem de quatro constatações: (i) as empresas de internet são tecnicamente superiores ao aparato estatal e possuem as ferramentas necessárias não apenas para identificar conteúdos e usuários, mas também exercer controle sobre eles; (ii) as plataformas têm acesso ao que os usuários dizem e fazem, conhecimento necessário para exercer o controle buscado pelo Estado; (iii) existem incentivos mútuos para essa relação simbiótica: o Estado economiza esforços e recursos cooptando as capacidades técnicas e informacionais das plataformas, ao passo que para as empresas precisam manter um bom relacionamento com o poder pública a fim de garantir a continuidade desembaraçada de seus serviços em um determinado país; e (iv) quanto mais poderosos e eficientes são os mecanismos de vigilância e controle das plataformas, mais valiosos eles passam a ser para o Estado, que por sua vez exerce pressão para que estas expandam ainda mais suas capacidades de vigilância, caracterizando um ciclo vicioso de poder e cooptação (BALKIN, 2018, p. 2010-2018).

controlados por empresas privadas (BALKIN, 2018, p. 2016). Esse esforço regulatório pode ser observado em diversas situações.

Na União Europeia, por exemplo, as grandes plataformas digitais firmaram um acordo com a Corte Europeia de Justiça (CEJ) a fim de desempenhar o que poderia ser considerada uma função administrativa: receber e julgar com base no direito europeu as reclamações de usuários relativas ao direito ao esquecimento. Com a decisão da CEJ, as empresas receberam a incumbência de desenvolver mecanismos próprios para receber e analisar demandas relativas a esse direito. O veredito também determinou que um solicitante somente poderia acionar a justiça após esgotar as vias privadas. Trata-se de uma dinâmica própria ao direito administrativo, em que se requer o esgotamento das instâncias internas para que seja possível recorrer à justiça comum. Essa determinação da CEJ caracteriza efetivamente a delegação de um serviço burocrático público a empresas privadas (BALKIN, 2018, p. 2016).

Outro exemplo paradigmático é a Lei para Melhorar a Aplicação da Lei em Redes Sociais (denominada Lei de Vigilância de Redes ou “NetzDG”), norma alemã que estabelece regras rígidas de moderação de conteúdo nas plataformas digitais. Aprovada pelo parlamento alemão em 2017, a nova legislação obriga as companhias a removerem conteúdo “manifestamente ilegal” de suas redes em até 24 horas após notificação dos usuários. O descumprimento da regra pode acarretar multas milionárias (BALKIN, 2018, p. 2010-2018).

No caso alemão, o juízo sobre o que seja manifestamente ilegal ou não, deve ser exercido pelas empresas tendo como base as leis do próprio país. Na prática, o Estado Alemão delegou às plataformas a competência de julgar e condenar os cidadãos que tenham publicado determinado tipo de conteúdo, contrário às suas leis locais. Nesse modelo, a plataforma atua como procurador, juiz, júri e executor, sem que usuários prejudicados sejam ouvidos e recebem as proteções do devido processo (BALKIN, 2018, p. 2010-2018).

Esses exemplos evidenciam a mudança no modo como o discurso e, por conseguinte, a liberdade de expressão, passaram a ser regulados na era digital. Enquanto o modelo dualista do século XX se restringia à relação entre o Estado e os emissores, a dinâmica da liberdade de expressão no século XXI demanda um modelo

distinto de regulação, que acomode os vários *players* que caracterizam a atual dinâmica de forças (BALKIN, 2018, p. 2010-2018).

2.3O TRIÂNGULO DE BALKIN: UM MODELO PLURALISTA DE REGULAÇÃO DA LIBERDADE DE EXPRESSÃO

Diante da nova realidade, Balkin propõe um modelo mais completo capaz de descrever a atual dinâmica mais adequadamente. O autor incorpora a antiga abordagem regulatória (“*old-school speech regulation*”) e a nova abordagem regulatória (“*new-school speech regulation*”). Juntas, as duas abordagens compõem o modelo pluralista de regulação da liberdade de expressão proposto pelo professor norte-americano (BALKIN, 2018, p. 2010-2018).

O modelo pluralista de regulação proposto por Balkin distingue-se de seu antecessor dualista por reconhecer a existência de um novo polo na dinâmica de funcionamento liberdade de expressão. Por essa razão, sua estrutura pode ser descrita na forma um triângulo (FIGURA 1), em que cada uma das pontas representa um dos sujeitos nessa nova dinâmica de forças (BALKIN, 2018, p. 2010-2018)

Na primeira ponta do triângulo está a figura dos “Estados-nação”, que designam as entidades políticas soberanas, com estrutura jurídica própria e que constituem um governo num dado território. É o caso do Estado brasileiro. Também se inserem nesse canto do triângulo os entes supranacionais de caráter público, como a União Europeia (BALKIN, 2018, p. 2010-2018)

A segunda ponta, por sua vez, abriga as companhias privadas detentoras da infraestrutura da internet. Tais corporações desempenham o papel crucial de garantir as condições de fala e comunicação aos indivíduos, oferecendo-lhes as ferramentas necessárias para tal serviço. Incluem-se aí as mídias sociais (ex.: Facebook, Twitter, TikTok etc.), os mecanismos de busca (também conhecidos como buscadores) (ex.: Google, Yahoo, Bing etc.), os provedores de internet (ex.: Claro, Vivo, Oi, etc.), e também os sistemas de pagamento eletrônico (ex.: PayPal, Wise, Mastercard, Visa etc.). Aqui está a principal inovação do modelo Balkin em relação ao seu antecessor de duas pontas (BALKIN, 2018, p. 2010-2018).

Além dessas duas figuras, há um terceiro grupo de sujeitos que ocupam a última ponta: os emissores (*speakers*). Na era digital, esse lugar acomoda não apenas os veículos da mídia tradicional (ex.: canais de televisão, jornais, revistas, editoras etc.), mas todo tipo de “locutor”, desde criadores de conteúdo, ativistas, até organizações da sociedade civil, hackers, trolls, entre outros (BALKIN, 2018, p. 2010-2018).

Para Balkin, a capacidade de comunicação no mundo digital hoje decorre da luta por poder entre os diversos polos que compõe esse diagrama. O jogo de forças entre as três pontas envolve não apenas o emprego, pelo Estado, de instrumentos regulatórios característicos das duas abordagens supracitadas (“old-school” e “new-school”), mas também a utilização, pelas próprias plataformas digitais, de mecanismos de regulação *privada*. Há ainda a pressão exercida tanto pela sociedade civil quanto pelo poder público junto às empresas para que estas moderem o fluxo de informações que corre por suas infraestruturas (BALKIN, 2018, p. 2010-2018).

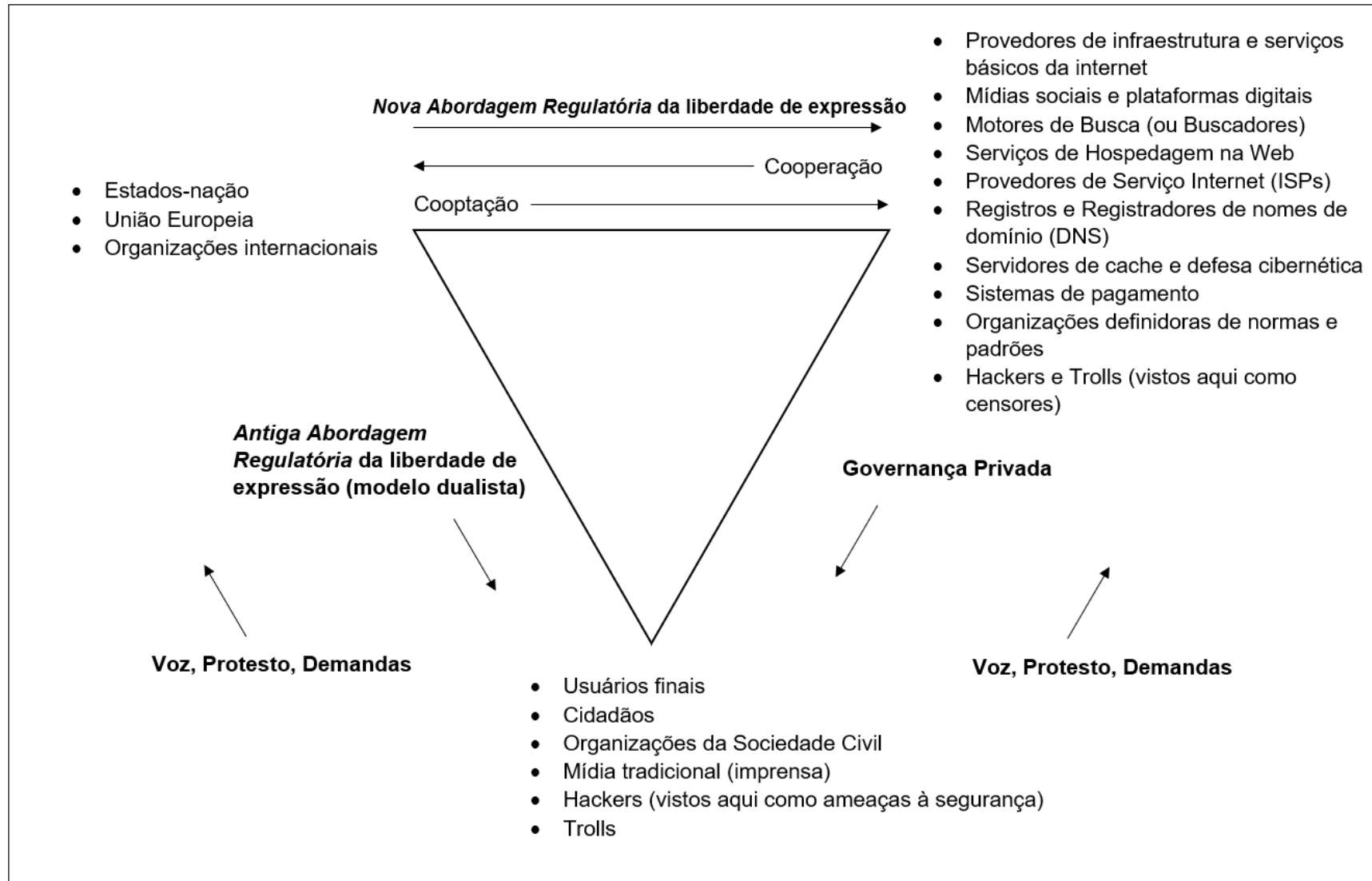


FIGURA 1: O Triângulo de Balkin (“Um modelo pluralista de regulação da liberdade de expressão”) (BALKIN, 2018, p. 2014, tradução do autor)

2.3.1 Problemas

Desse novo jogo de forças proposto por Balkin decorrem três problemas distintos. O primeiro deles é a ocorrência de censura colateral e censura prévia nas plataformas digitais. Essa prática resulta em grande medida da pressão exercida pelos governos pela remoção de certos conteúdos, por meio de mecanismos regulatórios típicos da *nova abordagem de regulação da liberdade de expressão*. (BALKIN, 2018, p. 2010-2018)

O segundo problema está no modelo de governança privada de conteúdo adotado pelas plataformas, marcado pela falta de transparência e de mecanismos de devido processo, o que resulta na ocorrência de arbitrariedades na moderação de conteúdo. (BALKIN, 2018, p. 2010-2018).

O terceiro problema decorre do modelo de negócios das grandes plataformas, baseado na coleta massiva de dados pessoais, e que torna os usuários finais vulneráveis à manipulação e à vigilância digital (BALKIN, 2018, p. 2010-2018).

Cada um desses problemas é detalhado adiante.

2.3.1.1 Censura prévia e colateral

A censura prévia, também conhecida como censura *ex ante*, designa um “controle preventivo das comunicações que se pretende veicular” (CABETTE, 2015). Nesta hipótese, um conteúdo é censurado antes mesmo de sua publicação.

Praticada abertamente pela ditadura militar no Brasil e por pelos estados absolutistas europeus dos séculos XVI e XVII, a censura prévia dava ao poder público a discricionariedade de vedar de antemão discursos, manifestações públicas, e outras formas de expressão. Para se expressar publicamente, exigia-se de certos sujeitos uma autorização administrativa ou judicial prévia (BALKIN, 2018, p. 2016-2019).

Conquanto o instituto da “censura prévia” seja vedado de forma quase absoluta no Estado Democrático de Direito, os mecanismos regulatórios da *nova abordagem* incentivam a ocorrência de um fenômeno semelhante. Ao constranger as empresas de infraestrutura da internet para que estas vigiem e removam certos conteúdos, o

Estado gera incentivos para que as plataformas pratiquem uma nova versão de censura prévia (BALKIN, 2018, p. 2016-2019).

É o que acontece, por exemplo, com a análise preliminar de conteúdo nas plataformas digitais, realizada por algoritmos logo após o “upload” de arquivos. Tais funcionalidades podem ser bastante úteis na identificação e remoção de conteúdos impróprios, como a pornografia infantil. No entanto, decisões automatizadas estão sujeitas ao erro e podem censurar indevidamente certos conteúdos, punindo usuários de forma abusiva. Isso tende a ocorrer com mais frequência se os processos de decisão são pouco transparentes e não incorporam garantias do devido processo, como o direito à informação e recurso (BALKIN, 2018, p. 2016-2019).

Há também a censura *ex post facto* ou posterior, na qual um conteúdo é removido logo após sua publicação, ou pouco tempo depois de ter sido divulgado. Nesse caso, há possibilidade de responsabilização civil ulterior da pessoa que produziu e/ou veiculou a comunicação considerada abusiva. Essa dinâmica pode fomentar o fenômeno da censura colateral (CABETTE, 2015).

A censura colateral ocorre quando um ente privado restringe o exercício da liberdade de expressão de alguém “*a fim de esquivar-se de eventuais responsabilidades civis ou criminais dispostas pelo poder público*” (CABETTE, 2015). É o que acontece, por exemplo, com a veiculação de conteúdo protegido por direitos autorais, ou de *revenge porn* na internet, práticas expressamente vedadas no ordenamento jurídico brasileiro. Caso os provedores de aplicações não removam esse tipo de material após tomarem conhecimento dele ou após a devida notificação judicial, poderão ser responsabilizados subsidiariamente pelas violações e danos decorrentes daquele conteúdo gerado por terceiros.²⁶ Trata-se, portanto, de uma modalidade indireta de censura (BALKIN, 2018, p. 2016-2019).

Sob a ameaça de tais punições, as empresas são incentivadas a exercer uma cautela excessiva na moderação de conteúdo de modo a evitar multas, processos e penalizações. Nessas circunstâncias, uma simples denúncia encaminhada por um único usuário, ou o uso de uma palavra-chave sinalizada como polêmica pelo algoritmo e que não tenha tido seu contexto analisado pela empresa, poderia ensejar

²⁶ Art. 19, § 2º e art. 21 do Marco Civil da Internet.

a remoção de um determinado conteúdo com base na lógica da mitigação dos riscos. Certamente, uma cultura de censura em excesso na moderação de conteúdo não é benéfica ao exercício desembaraçado da liberdade de expressão (BALKIN, 2018, p. 2016-2019).

A aplicação de penalidades e restrições à liberdade de expressão no mundo digital ocorre frequentemente sem transparência e garantias do devido processo. Muitas vezes, os usuários que tiveram seus conteúdos removidos, ou suas contas bloqueadas, não recebem a oportunidade de serem ouvidos. Também não há uma obrigação para que sejam apresentadas as razões da pena aplicada (BALKIN, 2018, p. 2016-2019).

2.3.1.2 Uma governança de conteúdo pouco transparente

Conforme aponta Kate Klonick (2018), as grandes plataformas de conteúdo constituem verdadeiras “governantes” do espaço virtual, posto que estabelecem relações reais de poder nos ecossistemas digitais, promulgando e aplicando regras de convivência por meio de estruturas complexas de governança privada.

Essas empresas não são meras facilitadoras da comunicação digital, ou simples intermediárias das relações sociais. Exercem enorme controle sobre os emissores (*speakers*) e comunidades que povoam seus domínios cibernéticos por meio de sistemas burocráticos intrincados.

Esses mecanismos robustos de governança interna, tem por objetivo criar condições mínimas de convivência em suas comunidades globais formada por uma enorme quantidade e variedade de usuários. Para impor um mínimo de ordem à multidão incontrolável e imprevisível de pessoas que utilizam seus espaços, utiliza-se uma combinação de contrato (ou seja, os termos de serviço e políticas de conteúdo) e algoritmos cada vez mais sofisticados (BALKIN, 2018, p. 2021-2032).

A fim de assegurar o cumprimento das regras da comunidade em escala global, as mídias digitais contam com moderadores humanos e algoritmos. Os funcionários humanos utilizam regras de fácil compreensão e aplicação para decidir um vasto número de controvérsias. Os casos mais difíceis de julgar são levados a instâncias superiores de decisão (BALKIN, 2018, p. 2021-2032).

Os algoritmos, por sua vez, constituem uma forma mais econômica de utilizar os recursos e ajudam na uniformização das decisões. Essas ferramentas podem substituir o julgamento humano em casos mais simples, e são especialmente úteis na análise prévia de conteúdo. No entanto, estão sujeitas a erros e imprecisões que podem gerar decisões arbitrárias (BALKIN, 2018, p. 2021-2032).

Um dos maiores problemas da governança privada de conteúdo é o caráter autocrático que tem guiado sua atuação, marcada por processos decisórios pouco transparentes e guiado por interesses privados. Isso fica evidente, por exemplo, no modo obscuro e pouco objetivo com que as políticas de conteúdo são implementadas, variando de acordo com a conveniência da situação e os interesses da empresa, conforme apontaram vazamentos de documentos internos do Facebook²⁷ em setembro de 2021 (BALKIN, 2018, p. 2021-2032).

Além disso, os usuários não contam com garantias processuais bem estabelecidas, como o direito a recorrer de sanções, ou mesmo de saber as razões pelas quais uma punição foi aplicada (BALKIN, 2018, p. 2021-2032).

Em que se pese os diversos problemas e dificuldades produzidos pela dinâmica privada de governança da liberdade de expressão, Balkin defende que o melhor caminho regulatório não passa pela adoção de regras excessivamente restritivas ou proibitivas à moderação privada de conteúdo. Também não se trata de um espaço de comunicação para o qual não se deva estabelecer limites. Muito pelo contrário. Caso fossem estabelecidas restrições proibitivas à moderação de conteúdo, é provável que estes espaços se tornassem ingovernáveis. Isso poderia diminuir significativamente seu valor para os usuários, e por conseguinte, para a sociedade (BALKIN, 2018, p. 2021-2032).

É fundamental que as plataformas digitais protejam a liberdade de expressão e reconheçam a importância de garantias do devido processo na resolução de litígios dentro da plataforma (BALKIN, 2018, p. 2021-2032).

²⁷ Em setembro de 2021, a ex-funcionária do Facebook Frances Haugen vazou documentos internos da empresa alegando, entre outras coisas, que a plataforma prioriza o "crescimento em detrimento da segurança", seja para proteger o funcionamento da democracia, quanto no cuidado com a saúde mental dos usuários. Os documentos também evidenciaram o modo diferenciado com que a empresa tratava certas figuras e personalidades poderosas ou influentes, que contavam com a boa vontade da rede social (BBC, 2021).

2.3.1.3 Vigilância e manipulação

Há que se pontuar, ainda, que esse vasto sistema de governança privada não tem como objetivo primeiro a proteção de um suposto “interesse público” ou a perseguição de um bem coletivo. Busca-se prioritariamente o interesse privado e a maximização dos lucros (BALKIN, 2018, p. 2021-2032).

Isso é bastante evidente no caso do Facebook. A companhia adotou regras de comunidade porque seu modelo de negócios pressupõe a existência um ambiente digital seguro, atraente e acolhedor para seus bilhões de usuários ao redor do planeta. O lucro da empresa está diretamente associado à sua capacidade de capturar e manter presa a atenção de seus usuário. Nesse cenário, um ambiente caótico e desgovernando poderia assustar as pessoas. Esse é um risco que a empresa (e seus acionistas) não podem correr (BALKIN, 2018, p. 2021-2032).

O modelo de negócios do Facebook requer que uma grande quantidade de indivíduos acesse regularmente a plataforma, consuma e crie conteúdo, a fim de que a empresa consiga vender o tempo e a atenção de seus usuários a anunciantes. A empresa oferece um serviço gratuito, e em troca o usuário lhe entrega sua atenção para receber anúncios personalizados (BALKIN, 2018, p. 2021-2032).

Nesse dinâmica, a vigilância digital desempenha um papel de grande importância. A fim de comercializar a atenção dos usuários a anunciantes, é necessário antes conhecer seus gostos, interesses e preferências, a fim de que o dinheiro investido pelos anunciantes não seja desperdiçado. Isso ocorre por meio da coleta de grandes quantidades de dados pessoais dos usuários (BALKIN, 2018, p. 2021-2032).

Há, portanto, um incentivo financeiro para que as empresas tornem seus conteúdos viciantes, conforme aponta Tim Wu (2016, p. 289–302). Isso porque quanto maior o tempo dispendido pelos usuários na plataforma, maior será a quantidade de anúncios apresentada, e maior o lucro da empresa. Esse tipo de negócio, portanto, explora a escassez de atenção dos usuários (BALKIN, 2018, p. 2021-2032).

2.3.2 Soluções propostas

A fim de solucionar esses problemas, Balkin (2018) propõe que sejam utilizadas diferentes soluções técnicas, regulatórias e administrativas para lidar com as relações do triângulo (BALKIN, 2018, p. 2032-2040).

Inicialmente, as práticas de censura prévia e colateral devem ser combatidas com uma regulação estrutural que fomente a competição, de modo a maior neutralidade, e prevenir a discriminação por conteúdo em sistemas de pagamento e serviços básicos da internet. Deve-se permitir o acesso e a participação de todos (BALKIN, 2018, p. 2032-2040).

Em segundo lugar, para enfrentar o problema das decisões arbitrárias e pouco transparentes, Balkin recomenda a adoção de garantias do devido processo na moderação de conteúdo realizada pelas plataformas, o que ele chama de “devido processo curatorial” (BALKIN, 2018, p. 2032-2040).

Em terceiro lugar, propõe-se que as grandes plataformas de conteúdo sejam consideradas “fiduciárias da informação” com deveres de confiabilidade e boa-fé para com os seus usuários (BALKIN, 2018, p. 2032-2040).

Para o autor, qualquer mecanismo regulatório que vise proteger a liberdade de expressão deve ter dois objetivos principais. O primeiro deles é prevenir, aperfeiçoar e limitar ao máximo a ocorrência de censura colateral e novas formas de censura prévia. O segundo objetivo é proteger os usuários contra novas formas de vigilância e manipulação (BALKIN, 2018, p. 2032-2040).

Estas soluções são aprofundadas a seguir.

2.3.2.1 Regulações pró-competição

Balkin defende a implementação reformas estruturais e políticas públicas que tenham como objetivo fomentar a competição. Conquanto possam ser considerados métodos típicos da *nova abordagem regulatória*, regulações que sirvam para garantir a competitividade não geram problemas como a censura ou regras arbitrárias de moderação de conteúdo (BALKIN, 2018, p. 2032-2040).

Dentre os exemplos de instrumentos regulatórios criados para este fim estão os padrões abertos²⁸ e as regras de neutralidade de rede. Também inserem-se aí leis antitruste capazes de lidar com o desafio dos oligopólios tecnológicos, que pelo seu agigantamento já tem sido denominados “datapólios” (STUCKLE, 2018, p. 278).

Corroborando a sugestão de Balkin, a proposta regulatória desenvolvida pelo *Observatorio Latinoamericano de Regulación de Medios y Convergencia* (PALLERO et al, 2020) pontua que:

“A regulação privada na Internet é produzida e agravada por um contexto de forte concentração de poder em poucas corporações multinacionais. A regulação pública sobre as atividades das plataformas deveria adotar medidas antimonopólio para enfrentar o cenário de concentração e ausência de concorrência, mas elas não estão incluídas nesta oportunidade. A simples ideia de que os principais espaços públicos para a circulação de informações e opiniões são todos controlados pela mesma empresa deveria forçar os órgãos antimonopólio a agir.” (PALLERO et al, 2020, p .8-9)

Parte do problema atual está no fato de que as plataformas foram aos poucos comprando seus competidores e consolidando sua posição dominante. Quanto maior a concorrência nos ecossistemas digitais, maior a probabilidade de surgirem inovações disruptivas, novas funcionalidades, melhores padrões de segurança e mais opções e modelos de governança dos espaços virtuais à disposição da sociedade (BALKIN, 2018, p. 2032-2040).

Com poucas alternativas às regras obscuras e às práticas arbitrárias de governança privada, os estímulos para a adoção de melhores práticas têm se mostrado insuficiente. Nesse sentido, uma maior variedade de redes sociais que ofereçam diferentes valores e estruturas de governança pode ajudar a aperfeiçoar o exercício da liberdade de expressão na internet (BALKIN, 2018, p. 2032-2040).

²⁸ Padrões abertos são aqueles “[...] publicamente disponíveis e não são controlados por nenhum governo ou corporação, tornam possível que quaisquer empresas, cidadãos e países se conectem e troquem informações com autonomia.” (CHEDE, 2008).

2.3.2.2 As responsabilidades da infraestrutura da internet

Balkin defende que os diferentes atores que compõe a infraestrutura privada da internet comportam diferentes responsabilidades na proteção da liberdade de expressão online. Nessa miríade de entes privados, é possível distinguir três grupos principais: (i) os serviços básicos de internet; (ii) os serviços de pagamento; e (iii) os curadores de conteúdo. Tendo em vista o escopo do trabalho, serão discutidas apenas as responsabilidades do terceiro grupo (BALKIN, 2018, p. 2032-2040).

A categoria dos curadores de conteúdo abrange tanto as mídias digitais (ex.: Facebook, Instagram, Twitter, TikTok, YouTube etc.) quanto os motores de busca, também conhecidos como buscadores (ex.: Google, Yandex, DuckDuckGo etc.). Essas empresas tomam decisões constantes e ubíquas sobre conteúdo, caracterizando um verdadeiro trabalho de curadoria (BALKIN, 2018, p. 2032-2040).

A tarefa de curadoria está na base dos serviços prestados pelas plataformas. Com base em regras pouco transparentes, as redes sociais, por exemplo, decidem a ordem na qual as publicações aparecerão no *feed* de um usuário. O Google, por sua vez, seleciona os links acha mais interessante com base no perfil e nos dados que possui sobre seus próprios usuários (BALKIN, 2018, p. 2032-2040).

Por atuarem como curadores e personalizadores de conteúdo, as plataformas não podem se esquivar da tomada de decisões sobre conteúdo. Por essa razão, possuem obrigações distintas para com sua membresia. Em essência, espera-se que os curadores sejam: (i) confiáveis na execução de seus serviços, e (ii) não-arbitrários na governança de suas próprias comunidades (BALKIN, 2018, p. 2032-2040).

2.3.2.3 As obrigações de confiabilidade, ética e profissionalismo

Os curadores digitais possibilitam o exercício da liberdade de expressão e a circulação de informações na era digital, e, portanto, exercem um papel civilizacional que transcende uma mera função de intermediário. Constituem, apesar dos problemas, guardiões da esfera pública e prestam um serviço imprescindível ao adequado funcionamento das estruturas democráticas (BALKIN, 2018, p. 2040-2047).

Em vista disso, os curadores digitais detêm obrigações morais e sociais para com o público. Esse dever decorre de três atividades de interesse público desenvolvidas pelas plataformas, afirma Balkin:

“Em primeiro lugar, eles **facilitam a participação pública** na arte, na política e na cultura. Em segundo lugar, eles **organizam as conversas públicas** para que as pessoas possam se encontrar e se comunicar facilmente. Terceiro, eles **fazem a curadoria da opinião pública** fornecendo *feeds* individualizados e resultados de buscas, além de reforçar normas de civilidade por meio das obrigações previstas em seus termos de serviço e diretrizes da comunidade.” (BALKIN, 2018, p. 2041, grifo do autor)²⁹

Conquanto estas não caracterizem obrigações legais, essas atividades trazem consigo responsabilidades inescapáveis àqueles que as realizam. Dessas responsabilidades surge a necessidade de que sejam elaboradas e aplicadas normas éticas e profissionais para a governança de conteúdo (BALKIN, 2018, p. 2040-2047).

As próprias plataformas de conteúdo reconhecem abertamente a função de relevância pública que exercem e reivindicam a necessidade de adoção de uma conduta pautada pela ética e pelo profissionalismo. Dentre os objetivos públicos declarados a garantia do “acesso à informação”, a “liberdade de expressão” e a “construção de comunidades” (BALKIN, 2018, p. 2040-2047).

Mais do que isso, essas empresas dizem fundamentar suas decisões de curadoria e moderação de conteúdo no dever de profissionalismo e na promoção do interesse público e social. Esses valores também são evocados para justificar a aplicação das regras da comunidade (BALKIN, 2018, p. 2040-2047).

Apesar disso, a opinião pública é uníssona em reconhecer que as plataformas digitais não têm incorporado espírito público que dizem defender. Com isso, passa a existir uma crescente demanda da sociedade para que esses entes privados se adequem a esses padrões éticos (BALKIN, 2018, p. 2040-2047).

Algo similar ocorreu com o jornalismo no início do século XX, período que a testemunhou escândalos envolvendo veículos de comunicação que deixaram de lado

²⁹ Original: “*First, they facilitate public participation in art, politics, and culture. Second, they organize public conversation so that people can easily find and communicate with each other. Third, they curate public opinion by providing individualized feeds and search results, and by enforcing civility norms through their terms-of-service obligations and community guidelines.*” (BALKIN, 2018, p. 2041)

sua função pública. Episódios de instrumentalização da imprensa para fins de propaganda ideológica, e a veiculação de anúncios que recorriam a práticas antiéticas para vender produtos, marcaram a época e ajudaram a moldar a opinião pública para que se estabelecessem normas e padrões éticos de profissionalismo na imprensa. Com o tempo, a mídia tradicional aceitou as obrigações profissionais para com o público inerentes ao seu ofício, que se refletiam no modo como as notícias eram reportadas (BALKIN, 2018, p. 2040-2047).

Nota-se, portanto, que o estabelecimento de normas profissionais no jornalismo moderno foi a culminação de um processo gradual. De forma semelhante, é possível observar uma curva de aprendizado no caso das mídias digitais. Até recentemente, as plataformas digitais se enxergavam como meras empresas de tecnologia. Hoje, contudo, têm sido confrontadas com suas responsabilidades públicas (BALKIN, 2018, p. 2040-2047).

Ao fim e ao cabo, a adoção de parâmetros éticos no jornalismo moderno mostrou-se de grande importância para o funcionamento das democracias. Essa valiosa conquista contribuiu para equilibrar o jogo de forças na sociedade, e contrabalancear o poder do Estado (BALKIN, 2018, p. 2040-2047).

Por essa razão, Balkin defende que é papel da sociedade, e não do Estado, exigir a adoção de normas profissionais aos curadores digitais. As plataformas podem cumprir essa função pública e prestar seus serviços com responsabilidade institucional, contribuindo para equilibrar o jogo de poderes em uma democracia (BALKIN, 2018, p. 2040-2047).

2.3.2.4 As obrigações legais do devido processo

Além disso, Balkin (2018, p. 2044-2047) defende que as plataformas devem sujeitar-se a obrigações típicas de devido processo. Essas obrigações distinguem-se, contudo, do dever de profissionalismo e ética também presente na atividade jornalística³⁰.

³⁰ Não há que se falar em garantias processuais no trabalho jornalístico de curadoria que a imprensa exerce sob seu próprio conteúdo. O editor de um jornal tem a liberdade de não publicar uma reportagem escrita por um de seus repórteres, ou mesmo de não veicular a carta de um leitor. Tal escolha além de

Uma vez que as plataformas exercem controle direto sobre o conteúdo criado por usuários, incluindo a imposição de medidas como a censura, e a imposição de restrições ao acesso e à utilização das plataformas, essas empresas passam a ter obrigações de devido processo para com seus usuários finais. Especialmente porque esse tipo de sanção tem o potencial de afetar significativamente o exercício da liberdade de expressão de um indivíduo (BALKIN, 2018, p. 2044-2047).

A moderação do conteúdo gerado por bilhões de pessoas traz consigo um interesse público que deve impor um dever de confiabilidade das mídias digitais para com os seus usuários. Por essa razão, é mister que as empresas forneçam garantias de que, quando censurarem ou limitarem a participação de um usuário, não incorram em excessos ou arbitrariedades (BALKIN, 2018, p. 2044-2047).

Os Princípios de Manila sobre Responsabilidade dos Intermediários delineiam algumas dessas garantias processuais. Dentre as recomendações estão o direito à explicação das decisões, o direito ao recurso, e a necessidade de clareza e acessibilidade para as regras e políticas de moderação de conteúdo (BALKIN, 2018, p. 2044-2047).

Balkin (2018) denomina esse conjunto de regras procedimentais dispostas nos Princípios de Manila e outros documentos similares como “obrigações do devido processo curatorial”. Para o professor estadunidense, o estabelecimento desse tipo de obrigação em lei não caracteriza necessariamente uma violação a autonomia privada ou à liberdade de expressão (BALKIN, 2018, p. 2044-2047).

Um caminho possível sugerido pelo autor é condicionar as imunidades legais contra a responsabilização por conteúdo de terceiro ao cumprimento dessas obrigações do devido processo. Isso faria com que o fornecimento desse tipo de garantia procedimental se tornasse um “portos seguro” contra a responsabilização de intermediários (BALKIN, 2018, p. 2044-2047).

Conforme delineado anteriormente, a Seção 230 nos Estados Unidos protege os moderadores de conteúdo da responsabilização civil por danos causados por conteúdo de terceiros. Tal provisão legal, em especial a cláusula do bom samaritano,

não configurar uma violação à liberdade de expressão, é exatamente aquilo se espera de um profissional à frente de um veículo da imprensa. Isso porque as organizações midiáticas e seu corpo editorial se responsabilizam diretamente pelo conteúdo que elas próprias publicam (BALKIN, 2018, p. 2040-2047).

garante nos EUA que as plataformas saiam “ílesas” mesmo editando, moderando, removendo, e censurando o conteúdo de seus usuários. Balkin (2018) questiona: essa ampla salvaguarda é razoável diante do enorme poder e controle exercido hoje por essas corporações?

Para o autor, a manutenção da imunidade dos provedores contra responsabilização de conteúdo alheio, deve ser mantida desde que estes aceitem adotar mecanismos procedimentais que garantam o devido processo na moderação de conteúdo. Caso contrário, perderão a ampla guarda legal concedida pela legislação (BALKIN, 2018, p. 2044-2047).

Esta proposta regulatória pode ser caracterizada como da *nova abordagem de regulação da liberdade de expressão*, posto que condiciona a imunidade de responsabilização por conteúdos de terceiros a um outra obrigação legal. A diferença, neste caso, é que, ao contrário do que normalmente ocorre, esse mecanismo será usado para proteger a liberdade de expressão e o interesse dos usuários (BALKIN, 2018, p. 2044-2047).

2.3.2.5 As obrigações legais dos fiduciários da informação

As mídias digitais do século XXI desenvolveram técnicas avançadas de vigilância, manipulação e controle de seus usuários. Por um lado, as plataformas possibilitam a participação em massa da população nos processos comunicativos, oferecendo-lhes seus serviços a preço zero. Por outro, promovem a coleta massiva de dados pessoais, em um ciclo vicioso de controle e vigilância digital. Tudo isso com o fim de capturar a escassa atenção dos usuários, e maximizar os lucros (BALKIN, 2018, p. 2047-2054).

Conquanto parte significativa dos veículos midiáticos continuem a se financiar por meio de anúncios, esses recursos tecnológicos não estão à disposição da mídia tradicional. O modelo de monetização introduzido pelas novas mídias digitais elevou o marketing a um novo patamar de alcance, personalização e lucratividade (BALKIN, 2018, p. 2047-2054).

Com base na intensa coleta de dados sobre seus usuários, as empresas criam perfis sobre cada indivíduo, classificando-os com base em gostos, interesses e

características pessoais. Esse processo, conhecido como perfilização, permite às plataformas oferecer anúncios ajustados a cada pessoa, e são, portanto, mais capazes de atrair a atenção dos usuários e influenciar seus comportamentos. Para os anunciantes, este é um modelo bem mais eficaz de alocação de recursos (BALKIN, 2018, p. 2047-2054).

Convém sublinhar que a moderação ou curadoria de conteúdo não envolve apenas a personalização de *feeds* ou resultados de busca para seus usuários finais. Trata-se de um processo mais amplo que inclui o uso de informações do próprio indivíduo a fim “controlar, moldar e governar” seus comportamentos. Ou seja, os curadores digitais são governantes não apenas no sentido de criar e aplicar regras de comunidade, mas também por sua tentativa de controlar e direcionar os usuários para satisfazer seus interesses econômicos utilizando para isso a vigilância (BALKIN, 2018, p. 2047-2054).

Esse enorme poder de controle e manipulação traz consigo riscos e perigos muito maiores, se comparado com aquele exercido pela mídia tradicional. O uso desse ferramental com o objetivo de manipular os usuários incorre também em uma grave violação do dever de confiança da empresa para com os usuários (BALKIN, 2018, p. 2047-2054).

Balkin argumenta que a enorme assimetria de forças e informações entre o usuário e a plataforma faz com que os indivíduos desenvolvam uma relação de dependência e vulnerabilidade com as empresas. Isso porque, por um lado, as mídias digitais conhecem muito acerca de seus usuários (por deles coletarem uma grande quantidade de dados pessoais), e usam essas informações para influenciá-los. Por outro, os usuários conhecem muito pouco ou nada sobre essas empresas e sua falta de transparência (BALKIN, 2018, p. 2047-2054).

É dessa relação de vulnerabilidade e dependência que surge um novo tipo de obrigação jurídica: as mídias digitais devem ser tratadas como *fiduciárias da informação* para com seus usuários. Justamente por conhecerem tanto quem utiliza seus serviços, e serem capazes de exercer tamanho controle e influência, é importante que haja um dever confiança para proteger a população contra abusos e manipulações (BALKIN, 2018, p. 2047-2054).

Nessa toada, é possível traçar um paralelo entre o dever de confiabilidade dos novos fiduciários da informação com a relação de confidencialidade entre médico e paciente, ou mesmo o sigilo profissional entre um advogado e seu cliente. Tais relações recebem proteção jurídica especial e estão previstas nos códigos de ética das respectivas categorias. O próprio art. 154 do Código Penal tipifica a violação de segredo profissional, punindo quem “*revelar [...], sem justa causa, segredo, de que tem ciência em razão de função, ministério, ofício ou profissão, e cuja revelação possa produzir dano a outrem*” (BALKIN, 2018, p. 2047-2054).

Tanto o sigilo médico quanto o segredo profissional na advocacia são previstos em lei, pois há o dever de que as informações sensíveis compartilhadas com esses profissionais não sejam usadas em prejuízo do cliente ou paciente. Antes, porém, espera-se que esses profissionais ajam com respeito e a confiabilidade devidas. Nesses casos, o ordenamento jurídico trata-os como “fiducitários” (BALKIN, 2018, p. 2047-2054).

Essas relações de fidúcia profissional pressupõe o dever de lealdade e a boa-fé para com as pessoas em uma relação de vulnerabilidade. Os fiduciários têm, portanto, três deveres principais: o dever de cuidado, o dever de confidencialidade, e o dever de lealdade para com seus clientes (BALKIN, 2018, p. 2047-2054).

Conquanto constituam relações jurídicas semelhantes, é importante estabelecer algumas distinções entre os fiduciários da informação e os deveres de confidencialidade no âmbito das relações profissionais citadas. Isso porque a natureza do dever de fidúcia varia de acordo com a natureza, com o papel social e o tipo de negócio (BALKIN, 2018, p. 2047-2054).

Por essa lógica, é possível concluir que as empresas digitais oferecem um serviço distinto daquele prestado pelos médicos ou advogados, o que dá origem a diferentes obrigações. Enquanto médicos e advogados são profissionais treinados e são remunerados diretamente por seus serviços, as novas mídias digitais e os buscadores oferecem serviços gratuitos para, em troca, coletar os dados pessoais de seus usuários, e utilizá-los para oferecer-lhes anúncios. Há, portanto, diferentes expectativas por parte dos consumidores (BALKIN, 2018, p. 2047-2054).

Trata-se de um relação ainda mais delicada, uma vez que a oferta de serviços gratuitos em troca de vigilância pode gerar conflitos de interesses entre as empresas

e seus usuários. Isso porque as companhias digitais têm incentivos para usar essas informações de formas que sacrifiquem o bem-estar de seus usuários, em benefício próprio. Alguns autores argumentam inclusive que dado esse potencial conflito de interesses, esse modelo de negócios deveria ser proibido. Na prática, deve-se analisar a natureza de um determinado negócio, e as legítimas expectativas dos consumidores para determinar se houve ou não conflito de interesses (BALKIN, 2018, p. 2047-2054).

Há, portanto, três deveres a serem cumpridos pelos fiduciários da informação: (i) o dever de confidencialidade, (ii) o dever do cuidado, e (iii) o dever de lealdade (BALKIN, 2018, p. 2047-2054).

As duas primeiras obrigações implicam que as plataformas devem guardar e proteger os dados de seus usuários de forma confidencial e segura. O dever de lealdade, por sua vez, se traduz no compromisso da empresa em não romper com as expectativas de seus clientes para benefício próprio, ou para o prejuízo do próprio usuário. Em outras palavras, é a responsabilidade de não enganar seus membros, ou trazer a eles desvantagem a fim de se beneficiar (BALKIN, 2018, p. 2047-2054).

2.4 AS ERAS REGULATÓRIAS DA GOVERNANÇA PRIVADA DE CONTEÚDO ONLINE

Do início dos anos noventa, com o surgimento da internet, até o segundo decênio do século XXI, a percepção da sociedade sobre a governança de conteúdo online atravessou intensas transformações. Essas mudanças de “humor” em relação à internet (de uma perspectiva positiva para outra negativa) deram origem a diferentes abordagens regulatórias em relação às plataformas digitais. Cada uma dessas abordagens representou um conjunto de valores e compreensões a respeito do uso das tecnologias, e trouxe ênfases distintas quanto à proteção ou limitação da liberdade de ação e expressão na internet (ZITTRAIN, 2019, p. 115).

Em um primeiro momento, com o desenvolvimento dos protocolos e padrões tecnológicos que permitiram o nascimento da internet, descortinou-se um enorme horizonte de possibilidades diante da humanidade. O despontar desse mundo digital trouxe consigo uma percepção idealizada da internet, que enfatizava seus efeitos

benéficos, e enxergava nela um espaço de conectividade, inovação e liberdade (ZITTRAIN, 2019, p. 115).

Até então, parte importante da opinião pública entendia que a internet não deveria estar sujeita a interferências ou controles externos que pudessem limitá-la. Um exemplo desse ideário é a Declaração de Independência do Ciberespaço, escrita pelo ativista John Perry Barlow (1996):

“Governos do Mundo Industrial, seus cansados gigantes de carne e aço, eu venho do Ciberespaço, a nova casa da Mente. Em nome do futuro, eu exijo a vocês do passado para nos deixar em paz. Vocês não são bem-vindos entre nós. Vocês não possuem autoridade soberana no lugar em que nos reunimos. [...] O Ciberespaço não se limita às suas fronteiras. [...] É uma força da natureza, e ela cresce através das nossas ações coletivas (BARLOW, 1996).”

Ocorre que, ao longo dos anos, essa visão utópica do espaço virtual, deu lugar à uma realidade bem menos promissora. A concentração de poder que então caracterizava o mundo off-line (“gigantes de aço”), passou a dominar também o mundo online. Agora os grandes conglomerados digitais, que caracterizam verdadeiros monopólios no ciberespaço, exercem seu domínio de forma desembaraçada, sem controles externos. Além disso, problemas sistêmicos como a desinformação, a disseminação de discursos de ódio, e a violência nas redes passaram a fazer parte do cotidiano dos usuários e trazer riscos a democracias contemporâneas (BOWERS & ZITTRAIN, 2020, p. 1-2).

A chegada das mídias digitais introduziu um novo modelo de criação e consumo de conteúdo em massa, assentado em uma cultura de “viralização” e potencialização de engajamento. Ao mesmo tempo, foram desenvolvidas e aperfeiçoadas poderosas ferramentas de marketing digital, capazes de direcionar anúncios personalizados e maximizar o efeito psicológico da mensagem veiculada. Isso transformou as plataformas digitais em alvos preferenciais para quem tinha como objetivo manipular, controlar e dissimular grupos e populações (BOWERS & ZITTRAIN, 2020, p. 1-2).

Não demorou muito para que essa combinação explosiva desse origem a inúmeros episódios dramáticos de interferências em eleições, e violência real que resultaram em mortes e até mesmo genocídio, como foi o caso de Myanmar (BOWERS & ZITTRAIN, 2020, p. 1-2).

Partindo de um olhar retrospectivo, Bowers e Zittrain (2020) identificam três momentos da governança das plataformas digitais: (i) a Era dos Direitos, (ii) a Era da Saúde Pública, e (iii) a era do Processo.

A Era dos Direitos, que se estendeu da década de noventa até os anos dez do novo milênio, foi caracterizada pelos esforços regulatórios em proteger esse recém-criado espaço virtual da interferência externa, e resguardar a esfera de liberdade dos usuários (BOWERS & ZITTRAIN, 2020, p. 1-2).

A Era da Saúde Pública, por sua vez, que se projeta até os dias de hoje, inverte essa lógica e dá lugar à pressão de reguladores para que as plataformas tomem decisões substanciais sobre conteúdo, estabelecendo formas de responsabilização para que as empresas de tecnologia enfrentem as mazelas digitais. Busca-se, portanto, “sanear” os espaços virtuais por meio de intervenções sistêmicas, mesmo que para isso seja necessário restringir a liberdade de expressão. Por demandar o controle e remoção de conteúdo problemático, a opinião pública se divide entre aqueles que consideram a moderação de conteúdo excessiva, e outros que a julgam insuficiente (BOWERS & ZITTRAIN, 2020, p. 1-2).

Por fim, os autores delineiam o surgimento de uma Era do Processo, que busca superar o dilema entre ampliar ou limitar a liberdade de expressão no espaço virtual, por meio do estabelecimento de processos de governança de conteúdo capazes de gerar consenso sobre o *modo* como devem ser tomadas decisões relevantes. Assim, o foco dos reguladores passa a ser no “como” ocorre o processo decisório de moderação, e não tanto sobre o que é ou não permitido. Busca-se, aqui, amalgamar consenso e legitimidade ao redor de procedimentais na governança de conteúdo (BOWERS & ZITTRAIN, 2020, p. 1-2).

Esse esquema de fases é útil para compreender os caminhos regulatórios percorridos até aqui, e traçar possíveis soluções regulatórias para os dilemas aparentemente inconciliáveis enfrentados até então.

2.4.1 A Era dos Direitos

Com o *boom* da internet, foi dada a largada em uma corrida de sucessivas inovações tecnológicas que revolucionaram os mais diversos setores da sociedade.

Em pouco tempo, surgiram inúmeras aplicações conectadas à rede mundial de computadores como serviços de e-mail, programas de mensagens instantâneas, serviços de chamada de áudio (VoIP), além de fóruns de discussão, e blogs (ZITTRAIN, 2019, p. 115).

Por constituir um ecossistema extremamente inovador e disruptivo, havia na época um enorme temor de que interferências governamentais pudessem atrapalhar o desenvolvimento desse novo espaço. Isso porque certos setores tradicionais (e poderosos) da economia, como as empresas de telecomunicação, começavam a ser afetados pela competição com a internet. Além disso, temia-se que aquela nova tecnologia fosse objeto de vigilância e controle estatal (ZITTRAIN, 2019, p. 115).

Havia, portanto, uma percepção geral de que era necessário proteger os espaços virtuais e seus atores contra ingerências externas, garantindo um ambiente de livre expressão e ação. (BOWERS & ZITTRAIN, 2020, p. 1-2). Essa compreensão era também enraizada na cultura jurídica estadunidense, que adota uma amplíssima aceção³¹ da liberdade de expressão, com base em interpretações da Primeira Emenda da Constituição norte-americana.

Foi nesse ‘espírito’ que nasceu o principal marco regulatório sobre a responsabilização de plataformas digitais nos EUA em 1996, a Seção 230 da Lei de Decência das Comunicações (*Communications Decency Act*, ou CDA). Tal provisão concedeu imunidade às plataformas de internet pelo conteúdo gerado por seus usuários. Para além disso, o dispositivo permitiu expressamente a moderação de conteúdo na internet, a fim de fomentar um espaço de convivência virtual saudável.

Este dispositivo paradigmático blindou as ainda jovens plataformas digitais, garantindo-lhes ampla proteção contra a responsabilização civil pelo conteúdo gerado por terceiros publicado em suas comunidades digitais. Mais do que isso, permitiu-lhes moderar e remover conteúdo por conta própria, sem temores de responsabilização. Essa provisão moldou o modo como a internet se desenvolveu, em especial porque

³¹ A cultura jurídica norte-americana sobre a liberdade de expressão é alicerçada nas interpretações dadas à Primeira Emenda da Constituição dos Estados Unidos, adotada em 1791, que prevê: “Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof; or abridging the freedom of speech, or of the press; or the right of the people peaceably to assemble, and to petition the Government for a redress of grievances”. Se interpretada de forma literal, esta norma parece conceder proteção quase absoluta à liberdade de expressão (AL SUR, 2021, p. 16-17).

as principais empresas de internet do mundo nasceram e estão sediadas nos EUA, submetendo-se à sua jurisdição e influência (KOSEFF, 2019, p. 8).

Tanto a Seção 230 quanto o Art. 19 do MCI caracterizam tipicamente a visão prevalente na Era dos Direitos, visto que tem por objetivo resguardar a livre circulação de ideias nesses espaços, sem criar para estes um ônus capaz de inviabilizá-lo. Trata-se de um marco legal que consagra um modelo de governança de conteúdo eminentemente autorregulatório, pois estimula as próprias empresas a exercerem o controle sobre a liberdade de expressão (BOWERS & ZITTRAIN, 2020, p. 2).

Convém ressaltar, ainda, que o contexto para o qual a Seção 230 foi criada era bastante distinto do atual. Em meados da década de noventa, a internet ainda dava seus primeiros passos e os usuários publicavam conteúdo próprio em espaços como blogs, fóruns e seções de comentários (BOWERS & ZITTRAIN, 2020, p. 2).

Isso mudou com a chegada avassaladora das plataformas digitais e suas poderosas ferramentas de indexação de dados e busca de informações. De acordo com Bowers e Zittrain (2020):

“As plataformas introduziram um novo modelo de consumo de conteúdo construído ao redor de *feeds* de informações, no caso dos resultados de pesquisa do Google, ou dos *feeds* de notícias do Facebook. [...] Em pouco tempo, [...] até mesmo o conteúdo mais tosco gerado por usuários aparecia ao lado de artigos meticulosamente verificados por jornalistas profissionais (BOWERS & ZITTRAIN, 2020, p. 3, tradução do autor)”³²

Ou seja, o mesmo espaço utilizado para divulgar conteúdo verificado e de qualidade, passou a ser também usado para disseminar informações enganosas, capazes de gerar danos à sociedade. Graças à imunidade contra responsabilização de conteúdo por terceiros, as plataformas em que são disseminados esse tipo de discurso recebem ampla proteção legal (BOWERS & ZITTRAIN, 2020, p. 3).

Contudo, a partir do momento que os riscos de prejuízo ganharam concretude e se agigantaram aos olhos da sociedade, passou-se a demandar posturas mais

³² Original: “*The platforms introduced a new model of content consumption built around feeds of material, like Google Search results or a Facebook Newsfeed. The feed contents were ranked by algorithms, optimizing, at least initially, for user engagement above all else, and making little visible distinction between different forms of content. They quickly flattened the landscape of internet content, presenting even the most rough-hewn of UGC in a common stream and on par with meticulously fact-checked articles by professional journalists.* (BOWERS & ZITTRAIN, 2020, p. 3)”

firmes das plataformas, em contraposição à acepção ampla e irrestrita da liberdade de expressão que havia prevalecido até então (BOWERS & ZITTRAIN, 2020, p. 3).

2.4.2 A Era da Saúde Pública

Bowers e Zittrain (2020) consideram que a segunda era de governança de conteúdo na internet se iniciou nos anos dez, com a popularização das redes sociais, e permanece até os dias de hoje. Nesse novo momento regulatório, espera-se que as plataformas atuem ativamente para identificar e remover conteúdo nocivo de suas comunidades. Essa ênfase no exercício de controle sobre o discurso dos usuários representa uma forte reação aos valores e compreensões defendidos na primeira era.

À medida que os efeitos negativos decorrentes da veiculação de conteúdo danoso começaram a ser sentidos na prática, intensificou-se a demanda pública por respostas regulatórias. Isso porque certas condutas (como a divulgação de informações falsas com o objetivo de dissimular ou enganar, ou mesmo a incitação à violência), se praticadas de forma sistêmica, podem gerar sérios danos aos indivíduos e às instituições (BOWERS & ZITTRAIN, 2020, p. 3-4).

Não é por acaso que a utilização de algoritmos como armas de controle e manipulação produziu diversos episódios trágicos. Um exemplo dramático é o massacre da minoria islâmica *rohingya* no Myanmar em 2017, incitado em grande medida por publicações comprovadamente mentirosas no Facebook que conclamavam o ódio contra esse grupo étnico. No ocidente, diversos países relatam interferências externas em seus processos eleitorais. Ao longo dos anos, incidentes dessa natureza predispuseram a opinião pública a pressionar as plataformas por regras mais duras de moderação de conteúdo. Essa mudança no “humor regulatório” deu origem à era da saúde pública (BOWERS & ZITTRAIN, 2020, p. 3-4).

A lógica de quem advoga essa perspectiva regulatória, passa pelo sopesamento dos benefícios e malefícios sistêmicos trazidos pelo atual modelo de governança de conteúdo, e na defesa de mecanismos regulatórios que compensem os excessos. Essa abordagem de regulação prioriza a “saúde” ou o “bem-estar” do público, ainda que isso implique na imposição de limites à liberdade de expressão. Enquanto o modelo anterior buscava a não-interferência na esfera de liberdade dos

usuários, neste, há uma pressão regulatória para que os abusos sejam reprimidos³³ (ZITTRAIN, 2019, p. 115).

Ao fim e ao cabo, a ênfase na remoção e na responsabilização por conteúdo nocivo da internet, envolve necessariamente a tomada de decisões regulares sobre o mérito das publicações. Esse juízo substantivo, na maioria das vezes exercido pelas próprias plataformas com base em suas regras e interesses, nem sempre é preciso e pode gerar situações reais de violação da liberdade de expressão. Opiniões e posicionamentos políticos constituem uma verdadeira área cinzenta, podendo ser considerados nocivos ou não a depender do contexto, e dos valores de quem analisa.

Há, portanto, um conflito aparentemente inconciliável de valores e concepções de mundo quando se discute o que deve ou não ser proibido na internet. De um lado, estão aqueles radicalmente contrários a quase todo tipo de censura, que adotam uma aceção ampla de liberdade de expressão e se identificam com o conjunto de valores da era dos Direitos. Do outro, se encontram aqueles que defendem um controle mais intenso do conteúdo publicado na internet, e que buscam sopesar a liberdade de expressão com outros bens e valores jurídicos como a democracia, a vida, e honra, e a saúde de outrem.

Os primeiros consideram que a moderação de conteúdo é frequentemente excessiva e abusiva. Os segundos, defendem que o controle exercido pelas plataformas é ainda insuficiente.

2.4.3 A Era do Processo

Para superar esse impasse, Bowers e Zittrain (2020) propõe a construção de um consenso entre reguladores, usuários e plataformas ao redor do *processo* de tomada de decisões na governança de conteúdo. Defende-se, portanto, o desenvolvimento de mecanismos de governança (ex.: regras de moderação) cuja

³³ Também se discute o impacto das regras que isentam as plataformas da responsabilização civil, como a *Section 230*, nos efeitos nocivos da governança de conteúdo. Por um lado, alguns políticos argumentam que a blindagem contra a responsabilização permite que as plataformas moderem conteúdo com um viés político, o que acarretaria censura indevida e controle excessivo por parte das plataformas. Por outro lado, alguns entendem que as mídias sociais não moderam o suficiente, e que as normas que a isenção de responsabilidade constitui um desincentivo para que estas procurem ativamente remover conteúdo nocivo (LASLO, 2019).

legitimidade seja amplamente reconhecida pelos dois lados, que sejam capazes de solucionar conflitos entre os diferentes valores. Desse modo, ainda que uma decisão de remoção de conteúdo desagrade um dos lados, o processo que levou àquela decisão será reconhecido como legítimo (BOWERS & ZITTRAIN, 2020, p. 4-5).

Nessa perspectiva centrada no processo, busca-se conciliar a proteção à liberdade de expressão, valor caro à era do direito, com as preocupações típicas da era da saúde pública. Para tanto, é necessário adequar os processos decisórios do atual modelo de governança privada de conteúdo, que além de pouco transparente, é guiado por interesses privados (BOWERS & ZITTRAIN, 2020, p. 4-5).

Por priorizar a lucratividade de seus serviços, as empresas tendem a tratar questões caras ao interesse coletivo como um problema de relações públicas, voltado para melhorar a imagem da empresa e reduzir a pressão externa. Desafios como a desinformação em meio à uma crise de saúde pública (ex.: infodemia), ou mesmo durante um processo eleitoral são muitas vezes tratadas como uma questão de “serviço ao consumidor” (BOWERS & ZITTRAIN, 2020, p. 4-5).

Contudo, essa postura não é adequada para lidar com questões dessa natureza, que demandam um certo distanciamento para evitar conflitos de interesses. Ou seja, não é recomendável que certas decisões sejam tomadas por órgãos internos às empresas, sujeitos a cadeias hierárquicas e estruturas de incentivo que podem comprometer sua independência (BOWERS & ZITTRAIN, 2020, p. 4-5).

Bowers e Zittrain (2020) vislumbram dois conjuntos de soluções regulatórias processuais para o dilema da governança privada de conteúdo online: (i) estabelecer novos deveres às plataformas digitais, reconhecendo-as como “fiduciárias de conteúdo” e (ii) delegar a tomada de decisões relevantes sobre governança de conteúdo para um ente externo independente.

2.4.3.1 Soluções propostas

Em primeiro lugar, os autores argumentam que é preciso tratar as plataformas como “fiduciárias de conteúdo”. A relação assimétrica entre as plataformas e os usuários, coloca-os em uma posição de vulnerabilidade e desvantagem. Sob estas circunstâncias, o Estado pode impor o dever de confiança (a relação de

fiduciariade) da parte em vantagem para com aquela em posição inferior. Caso o fiduciário não cumpra os deveres de confiança que decorrem dessa relação, poderá ser punido (BOWERS & ZITTRAIN, 2020, p. 5).

Esta proposta é semelhante à feita por Balkin, que sugeriu tratar os provedores como “fiduciários da *informação*”. A distinção entre os dois autores, contudo, está no fato de que Balkin se concentra na proteção à privacidade dos usuários, enquanto Bowers e Zittrain focam na governança de conteúdo.

Assim como a proposta de Balkin, aqui os autores defendem que as plataformas devem assumir o dever de proteger e garantir os interesses de seus usuários nas dinâmicas de governança de conteúdo. Isso implica no desenvolvimento de padrões éticos e profissionais no exercício da curadoria de conteúdo de terceiros. Um exemplo é a adoção medidas do devido processo (BOWERS & ZITTRAIN, 2020, p. 5).

Por último, propõe-se que as decisões mais difíceis e relevantes sejam delegadas a entidades externas, responsáveis por definir padrões e regras gerais de moderação. Por possuírem o distanciamento adequado, entes dessa natureza podem emitir juízos de valores que protejam de forma justa o interesse e o bem-estar dos usuários, sem se preocupar com retaliações ou influências indevidas. É por essa razão que decisões-chave sobre questões fundamentais da governança de conteúdo devem ocorrer, ao menos em parte, fora das empresas e distantes da influência dos interesses privados (BOWERS & ZITTRAIN, 2020, p. 6-7).

2.5 UMA ABORDAGEM PROCEDIMENTAL PARA REGULAR AS GRANDES PLATAFORMAS DE CONTEÚDO: CONSENSOS, PROPOSTAS E DESAFIOS

As molduras conceituais desenvolvidas por Balkin, Bowers e Zittrain fornecem os elementos necessários para delinear uma proposta regulatória que enderece as disfuncionalidades da atual dinâmica de forças que caracteriza a liberdade de expressão na era digital. A partir das convergências identificadas na delimitação do problema e nas soluções propostas pelos autores de referência, buscar-se-á neste subcapítulo construir as linhas gerais de uma “perspectiva procedimental” para regular as plataformas digitais.

Trata-se de uma abordagem regulatória “procedimental” pois se concentra nos *procedimentos*³⁴ de tomada de decisão do atual modelo de governança privada de conteúdo online. Por um lado, propõe-se a adoção de garantias processuais típicas do devido processo na moderação de conteúdo pelas plataformas. Por outro, enfatiza-se a importância de um organismo independente de fiscalização e controle, capaz de contrabalancear o poder das grandes plataformas, atuando como um mecanismo de freios e contrapesos. É possível situar essa perspectiva no paradigma do “constitucionalismo digital”³⁵.

2.5.1 O reconhecimento de novos direitos e deveres

Tanto Balkin (2018) quanto Bowers e Zittrain (2020) argumentam que as enormes assimetrias de poder e informação existentes entre os usuários e as corporações de internet devem motivar o reconhecimento de um novo vínculo jurídico de confiabilidade (ou “fiduciariedade”) entre as duas partes. Dessa relação devem decorrer os deveres de cuidado, lealdade e boa-fé das plataformas para com os usuários, bem como o estabelecimento de proteções legais e obrigações profissionais, similares àquelas existentes entre médico e paciente, e entre advogado e cliente.

Seguindo esta mesma perspectiva, relatórios publicados pelos governos da França (2019) e do Reino Unido (2019) propõe o estabelecimento de um “dever de cuidado” da plataforma para com os usuários, bem como de uma “obrigação de transparência” aliada a uma “obrigação de defender a integridade dos usuários” (ARHEGAS *et al*, 2021, p. 19-21; AL SUR, 2021, p. 29-30).

³⁴ No contexto jurídico os termos “processo” e “procedimento” possuem significados distintos, sendo o primeiro mais abrangente e indicando “o instrumento para se conseguir a prestação jurisdicional, com uma sucessão de atos processuais específicos” e o segundo mais singular, entendido como “o modo pelo qual esses atos processuais devem ser cumpridos, ou seja, qual rito seguirão” (VALERIO, 2019). Para Daniel Assunção Neves (2018, p. 159), “procedimento é a exteriorização do processo”. No caso da governança de conteúdo online, embora possam ser utilizados em contextos semelhantes, o termo “procedimento” indica um ato mais específico na moderação de conteúdo, enquanto o vocábulo “processo” incorpora um sentido mais global, que inclui os macro-processos de decisão, como a formulação das normas de moderação, e mecanismos de controle externo das decisões.

³⁵ O termo “constitucionalismo digital” designa o uso das estruturas e mecanismos do constitucionalismo moderno para guiar a governança do espaço digital. Conquanto não haja um consenso entre doutrinadores a respeito de sua definição, a ideia de constitucionalismo digital designa uma derivação do constitucionalismo moderno, que busca “estabelecer e garantir a existência de um quadro normativo para a proteção dos direitos fundamentais e o equilíbrio de poderes no ambiente digital” (CELESTE, 2018, p. 20).

Ao tratar os provedores de aplicação de internet como “fiduciários” do conteúdo gerado por seus usuários, deve-se reconhecer, portanto, a existência de direitos destes perante aqueles. Propostas regulatórias como a Medida Provisória nº 1068 acertam nesse sentido, mas acabam ‘pecando’ pelo excesso, engessar a moderação de conteúdo com parâmetros proibitivos de análise e sem um amplo debate público prévio.

Dada a enorme relevância das grandes plataformas no debate público e na vida social do país, não é adequado enxergá-las como senhoras absolutas de “espaços virtuais privados”, nos quais ditam as regras sem dar nenhuma satisfação à sociedade. Se essas corporações constituem verdadeiros “governos” do ambiente cibernético, é natural pensar no estabelecimento de direitos dos “governados” face a seus “governantes”. Para isso, deve-se tratar como “expressão” o que as empresas frequentemente designam em termos de negócio como “conteúdo” (MONTEIRO *et al*, 2021, p. 25; KLONIC, 2018, p. 1598–1670)

2.5.2 Entre uma “regulação autoritária” e uma “autorregulação corporativa”

Há, no entanto, uma polarização entre um modelo de “autorregulação corporativa”, sem *accountability* ou mecanismos de controle, e um modelo de “regulação autoritária”, que intervém de forma exacerbada na governança de conteúdo online. Essa regulação indevida ocorre quando se estabelecem regras excessivamente proibitivas às plataformas, como no caso da MP 1068, ou pela imposição de obrigações aos provedores para que estes monitorem ostensivamente o conteúdo gerado por seus usuários, a fim de detectar supostas violações da lei ou impedir futuras ilicitudes, como ocorre com a NetzDG (PALLERO *et al*, 2020, p. 7).

É preciso encontrar um equilíbrio entre essas duas abordagens, com base na noção de que a internet deve permanecer como um ambiente livre, aberto, seguro e democrático. Conquanto a opinião pública tenha se insurgido³⁶ mais recentemente contra as grandes empresas de tecnologia, é inegável que estas desempenham um

³⁶ O recente fenômeno conhecido como “techlash” designa a “animosidade pública em relação às grandes empresas de tecnologia, especialmente as do Vale do Silício”. O termo é um “acrônimo formado a partir das palavras *technology* e *backlash* (forte reação negativa a questões sociais ou políticas)” (MENA, 2020).

papel de fundamental importância para toda a sociedade. Por mais complexa e problemática que seja, a moderação de conteúdo é uma tarefa necessária, e sensível para a proteção aos direitos fundamentais (MONTEIRO *et al*, 2021, p. 26-27).

Além disso, uma regulação balanceada das grandes plataformas não deve constituir um óbice à inovação, competição ou desenvolvimento de novos modelos de governança de conteúdo (PALLERO *et al*, 2020, p. 5). Um exemplo de inovação neste particular é o Reddit. A rede social combina uma política global própria para a moderação de conteúdo, com mecanismos de controle exercidos pelos próprios usuários em suas comunidades (*subreddits*), constituindo uma espécie de modelo “federalista” para a moderação de conteúdo (MONTEIRO *et al*, 2021, p. 26).

Conforme destacou Balkin (2018), é preciso evitar uma concentração ainda maior desse nicho, que já conta com fortes traços oligopolísticos observáveis nas posições dominantes dos *gatekeepers*.

2.5.3 O “devido processo” na moderação de conteúdo

Contra a censura prévia e colateral, e eventuais arbitrariedades praticadas pelo modelo de governança privada, Balkin (2018) propõe que sejam adotados mecanismos do “devido processo” na curadoria de conteúdo. Bowers e Zittrain (2020), por sua vez, apostam na construção de um amplo consenso ao redor de regras aplicáveis à moderação de conteúdo, capazes de superar o impasse entre a tutela da liberdade de expressão e a proteção a outros bens e valores jurídicos. Com isso, buscam amalgamar legitimidade ao processo de tomada de decisões sobre conteúdo, transcendendo controvérsias substanciais.

O estabelecimento de padrões típicos do devido processo possibilitam contrabalancear as enormes assimetrias entre plataformas e usuários, sem que isso inviabilize a atuação das empresas. O acesso a esse tipo de mecanismo permite que usuários que tenham sido tratados de forma injusta possam ser ouvidos, e tenham condições de exigir explicações contra decisões arbitrárias na moderação de conteúdo (BERGMAYER, 2018, p. 7).

Já existem recomendações bastante robustas para a adoção desses mecanismos, como os Princípios de Manila (2015), os Princípios de Santa Clara

(2018) e o documento “Padrões para uma regulação democrática das grandes plataformas” (2020), produzido em colaboração por diversas organizações latino-americanas. Com base nesses documentos, propõe-se diretrizes mínimas para a construção de um ambiente virtual em que os direitos e liberdades fundamentais possam coexistir.

2.5.3.1 Direito à informação e dever de transparência

Os usuários devem ter acesso a informações claras, precisas, didáticas e de fácil alcance sobre os termos de uso e as políticas de conteúdo adotadas pelas plataformas, assim como ao modo como estas regras são aplicadas. Trata-se, sobretudo, de uma obrigação de transparência em relação às regras de moderação e a forma com que estas são aplicadas.

É necessário garantir que os usuários sejam informados com especial destaque sobre situações que possam afetar o exercício de seus direitos. Em outras palavras, é preciso explicitar aos utilizadores quais circunstâncias podem resultar na imposição de sanções como a remoção de um determinado conteúdo, a exclusão de uma conta, ou desmonetização de determinada página.

Nesse diapasão, o acesso à informação deve alcançar precipuamente as regras que disciplinam a curadoria e a moderação de conteúdo. Isto é, as plataformas devem explicitar quais são os critérios utilizados pelos algoritmos a fim de ordenar, priorizar, reduzir o alcance de uma determinada publicação, inclusive indicando tais efeitos aos usuários afetados.

Desse modo, a transparência em relação às regras de moderação serve para dar aos usuários a previsibilidade necessária em relação a quais conteúdos podem vir a prejudicá-los. O documento “Padrões para uma regulação democrática das grandes plataformas que garanta a liberdade de expressão online e uma Internet livre e aberta” (2020), propõe que as plataformas exponham de forma acessível, clara e explícita:

“A. Que tipos de conteúdo e atividades estão proibidos em seus serviços.

B. Quais são os critérios e mecanismos de curadoria e moderação de conteúdo; quais são controlados diretamente pelo usuário e quais não são; como o algoritmo de curadoria de conteúdo utilizado afeta a visibilidade de conteúdos.

C. Em que casos, quando e como a automação de remoção de conteúdo se aplica.

D. Em que casos, quando e como a revisão humana de conteúdos se aplica. Em particular, que critérios foram utilizados para a tomada de decisões, levando em conta o contexto, a ampla variação de nuances idiomáticos e o significado e as peculiaridades linguísticas e culturais dos conteúdos sujeitos a uma possível restrição.

E. Quantos moderadores tem, descrevendo em detalhes seu perfil profissional (experiência, especialização ou conhecimento), sua localização geográfica e sua distribuição de tarefas (em termos de temas, áreas geográficas etc.).

F. Quais são os direitos dos usuários em relação ao conteúdo gerado e publicado por eles mesmos e as políticas aplicadas pela empresa a esse respeito.

G. Como são utilizados e qual é o tratamento dos dados pessoais dos usuários, incluídos os pessoais e sensíveis, feito para tomar decisões algorítmicas que impactem os seus direitos. (PALLERO *et al*, 2020, p. 16-17)”

2.5.3.2 Direito à notificação e à decisão fundamentada

O direito à notificação e à decisão fundamentada devem garantir que os provedores de redes sociais expliquem de forma clara e específica o motivo pelo qual um determinado conteúdo foi removido, limitado ou restringido, ou mesmo as razões pelas quais uma conta foi bloqueada, suspensa ou excluída.

A notificação deve, no mínimo, especificar a cláusula em relação a qual alega-se a violação do usuário. Também é importante indicar com precisão qual conteúdo foi impugnado, informando também o modo pelo qual ele foi detectado, analisado e contestado. Do mesmo modo, é necessário que a notificação contenha informações claras sobre como o usuário pode recorrer da decisão.

De acordo com os Princípios de Manila (2015), uma notificação sobre decisões que restrinjam conteúdos de usuários, devem conter, no mínimo:

- “1. As razões pelas quais o conteúdo em questão viola as políticas de restrição de conteúdos do intermediário;
2. O identificador de Internet e descrição da alegada violação das políticas de restrição de conteúdo;
3. Informações de contato da parte emissora da solicitação ou de seus representantes, a menos que isso seja proibido por lei;
4. Uma declaração de boa-fé indicando que as informações fornecidas estão corretas. (EFF, 2015, p. 3)”

2.5.3.3 Direito de defesa e de apelação

Os usuários devem ter garantido o direito de recorrer contra decisões de moderação de conteúdo que lhe digam respeito, tendo assegurado a ampla defesa e o contraditório. Isso significa que, caso a plataforma tome uma decisão que restrinja de forma ilegítima ou desproporcional o direito à liberdade de expressão, o usuário poderá apelar da medida, permitindo que esta seja reavaliada.

Dessa forma, será possível ponderar se a medida aplicada em um caso específico se adequou aos princípios de proporcionalidade, necessidade, idoneidade e progressividade na persecução de uma finalidade específica (CIDH, 2013, p. 55). Também devem ser considerados fatores como:

“[...] a gravidade dos danos, seu alcance, a recorrência das violações, e o impacto que tal restrição poderia ter na capacidade de Internet para garantir e promover a liberdade de expressão em relação aos benefícios que a restrição traria para a proteção de outros direitos (PALLERO *et al*, 2020, p. 18)”

Conquanto algumas plataformas de conteúdo já forneçam mecanismos de recurso contra decisões de moderação, esta é uma prática ainda muito insipiente.

As propostas apresentadas pelo *Observatorio Latinoamericano de Regulación de Medios y Convergencia* (PALLERO *et al*, 2020) indicam cinco recomendações a serem garantidas no exercício do direito de recurso dos usuários:

A. Os mecanismos de recurso devem ser muito acessíveis e fáceis de usar.

B. Os recursos devem estar sujeitos a revisão por uma pessoa ou painel de pessoas que não estiveram envolvidas na decisão inicial e que não sejam parte interessada.

C. Os usuários devem ter o direito de propor novas evidências ou materiais a serem considerados em sua opinião.

D. As apelações devem resultar em rápida determinação e resposta ao usuário.

E. Qualquer exceção às garantias de dupla instância consagradas nos padrões de devido processo deve ser claramente divulgada e compatível com os demais princípios internacionais de direitos humanos. (PALLERO *et al*, 2020, p. 21-22)”

Para além das instâncias recursais oferecidas pela própria plataformas de conteúdo, não devem existir restrições contratuais para que os usuários recorram a

outras vias legais para contestar medidas e buscar a reparação de danos. Por certo, a possibilidade de acionamento legal no próprio país do usuários deve ser assegurada. Cláusulas contratuais que visem impedir a possibilidade de contestação na jurisdição do utilizador exploram de forma abusiva a assimetria de forças entre as partes contratuais.

2.5.4 Um organismo de controle independente

No direito público, o princípio do devido processo incorpora também o direito a um tribunal imparcial, que constitua um terceiro desinteressado para julgar a causa (BERGMAYER, 2018, p. 31; MOSTERT, 2020, p. 21). Conforme enfatizado por Bowers e Zittrain (2020), decisões sensíveis de moderação de conteúdo, que possam trazer grandes repercussões no jogo democrático, ou que envolvam diretamente o interesse público, deveriam ser delegadas a um ente externo, alheio às influências das grandes corporações, bem como às pressões de políticos e outros atores. Essa instituição pode ser também responsável por definir padrões e regras gerais de moderação, de caráter recomendatório ou normativo.

Contudo, ainda não há uma resposta definitiva sobre qual seja o formato ideal dessa instituição. Diversas propostas têm surgido ao redor do mundo para tentar suprir essa lacuna. Na Alemanha, por exemplo, a NetzDG criou um “Comitê de Peritos” para sanar dúvidas quanto à definição de termos e responder solicitações das plataformas em relação casos duvidosos (AL SUR, 2020, p. 29).

A União Europeia, por sua vez, tem discutido a criação de uma “Coordenadoria de Serviços Digitais” responsável por responder a tais questões em cada país-membro. Esse novo órgão está previsto no *Digital Services Act*, proposta legislativa apresentada ao Parlamento Europeu e ao Conselho Europeu em dezembro de 2020, e seria construído nos moldes das autoridades de proteção de dados já existentes no bloco de países (MONTEIRO *et al*, 2021, p. 30).

No relatório “Online Harms White Paper” publicado pelo governo britânico em abril de 2019, defende-se a criação “de um novo sistema de regulação das redes sociais a ser implementado por um órgão regulador público e independente”, responsável por implementar “padrões para garantir a segurança dos usuários nas

redes sociais” e garantir “a proteção da liberdade de expressão no ambiente digital”. Esse ente também seria responsável por, entre outras coisas, produzir “códigos de boas práticas” para as redes sociais, e monitorar a implementação do “dever de cuidado” para com os usuários (ARCHEGAS *et al*, 2021, p. 19-21, *apud* UK, 2019). Esse documento fundamentou a propositura de um projeto de lei no parlamento britânico em maio de 2021, denominado “Online Safety Bill”. Nessa proposta legislativa, discute-se a delegação das atribuições supracitadas ao “OFCOM”, entidade que regula os serviços de telecomunicações naquele país (UK, 2021).

Esse mesma visão é compartilhada pelo governo francês, em relatório publicado em maio de 2019, que defende a constituição de um órgão independente, formado por representantes do governo francês, responsável e por monitorar a implementação dessas das obrigações de transparência de defesa da integridade dos usuários. Trata-se de um modelo que combina o estabelecimento de regras mínimas “para preservar o interesse público” com a autorregulação das empresas. (ARCHEGAS *et al*, 2021, p. 20-21, *apud* FRANÇA, 2019).

No Brasil, o Congresso Nacional tem se debruçado sobre essas questões no âmbito do Projeto de Lei nº 2.630 de 2020, popularmente conhecido como “PL das Fake News”. Contudo, ainda não se sabe ao certo qual seria a engenharia institucional mais adequada para endereçar esse tema. Qualquer iniciativa regulatória, deve, antes de mais nada, passar por um amplo debate público, trazendo todos as partes interessadas à discussão

2.5.5 Desafios e limitações

Conquanto a perspectiva procedimental tenha se mostrado um caminho promissor para enfrentar os problemas e disfuncionalidades da atual dinâmica de autorregulação corporativa, é preciso problematizá-la com outros pontos de vista a respeito do tema.

Evelyn Douek (2019, p. 8-11) argumenta que forçar as empresas de internet a adotar garantias de devido processo na moderação de conteúdo pode trazer uma série de dificuldades. A primeira delas é a escala. De acordo com dados divulgados em 2018, o Facebook conta com quase 2,27 bilhões de usuários em todo o mundo e

toma cerca de três milhões de decisões sobre moderação de conteúdo diariamente. Esse número de decisões supera com folga a quantidade de sentenças judiciais tomada em um único ano pela maioria dos sistemas judiciais do planeta (FACEBOOK, 2018; DOUEK, 2019). Além disso, o próprio fundador e CEO do Facebook, Mark Zuckerberg, reconheceu que pelo menos uma em cada dez decisões sobre moderação de conteúdo tomada em sua plataforma incorre em algum equívoco (ZUCKERBERG, 2018).

Diante desse cenário, Evelyn pontua que erros são inevitáveis dada a enorme escala e a velocidade com que as decisões de moderação de conteúdo são tomadas. Desse modo, a sociedade precisa escolher “quais tipos de erros são preferíveis” (“*trade-offs*”) e buscar equilibrar os elementos em cada caso, levando em consideração o tipo de conteúdo analisado. Soluções regulatórias procedimentais devem levar em consideração fatores como escalabilidade, custos envolvidos (inclusive os recursos humanos e tecnológicos à disposição), a qualidade das decisões (justiça substancial), sua velocidade (decisões rápidas podem ser necessárias em certas situações, mas também estão mais sujeitas à erro), e o impacto concorrencial naquele mercado específico (DOUEK, 2019, p. 8-11).

Além disso, há também dificuldades em adaptar para relações privadas noções do “devido processo legal” utilizadas pelo direito público para lidar com a assimetria de poder entre o Estado e os cidadãos. Ainda que a adoção de mecanismos processuais pelas plataformas possa ser benéfica (especialmente por agregar legitimidade), a pesquisadora australiana pontua a ideia de um “devido” processo não deve ser confundida com uma obrigação em fornecer um processo “perfeito” em cada caso. Isso porque, mesmo no direito processual público, existem diferentes tipos de processos, cada qual com distintos níveis de proteções legais, adequados a circunstâncias distintas, próprias às necessidades e características de cada caso. A escolha de qual processo é o mais adequado, envolve um juízo sobre quais são os interesses privados envolvidos, os ganhos adicionais de valor pela adição de uma nova garantia processual e a sobrecarga que essa adição traz. Portanto, é preciso tratar regras procedimentais de forma sistêmica, tendo sempre em vista o todo, e os objetivos gerais de um determinado serviço (DOUEK, 2019, p. 8-11).

É preciso reconhecer que diferentes tipos de conteúdo (ex.: *spam*, violência explícita, pornografia infantil, discursos de ódio, *bullying*, etc), devem ser tratados de forma distintas, com processos distintos. Enquanto alguns tipos de conteúdo podem ser moderados de forma mais automatizada e com menos garantias procedimentais (é provavelmente o caso do *spam*), outras categorias de conteúdo podem demandar uma análise mais detida, e mais garantias processuais (DOUEK, 2019, 8-11).

3 O PROJETO DE LEI Nº 2630/2020

O objetivo deste capítulo é eminentemente documental. Com ele, pretende-se expor os aspectos mais relevantes do PL 2.630/2020, também conhecido como “PL das Fake News”, concentrando-se nos dispositivos que dizem respeito ou se relacionem ao *processo* de moderação e governança de conteúdo das plataformas. Para fins de recorte metodológico, optou-se por utilizar o texto aprovado pelo Senado Federal e seu relatório, independentemente da tramitação que segue na Câmara dos Deputados.

Na primeira seção são apresentadas informações gerais e contextuais a respeito da proposta legislativa. Em seguida, parte-se para a exposição das provisões legislativas atinentes ao presente trabalho, a saber, as disposições preliminares relevantes (art. 1, 2, 3, 4 e 5), os procedimentos de moderação (art. 12), o Conselho de Transparência e Responsabilidade na Internet (art. 25, 26, 27, 28 e 29), o ente de autorregulação regulada (art. 30), as sanções (art. 31) e as disposições finais pertinentes (art. 31, 32, 33, 34, 35 e 36).

3.1 CONTEXTUALIZAÇÃO

Em 30/06/2020 o Plenário do Senado Federal aprovou o Projeto de Lei 2.630/2020 que “institui a Lei Brasileira de Liberdade, Responsabilidade e Transparência na Internet”. O texto inicial foi apresentado pelo senador Alessandro Vieira (Cidadania-SE) em 13 de maio de 2020 e tramitou na casa sem passar por comissões temáticas. Foram apresentadas 152 emendas, e quatro relatórios consecutivos antes da votação final. Ao fim e ao cabo, foi aprovado um texto alternativo, na forma de um substitutivo, elaborado pelo senador Angelo Coronel (PSD-BA), relator da proposta. A votação contou com 44 votos favoráveis, 32 contrários e 2 abstenções (BRASIL, 2020).

Em seguida, o projeto seguiu para Câmara dos Deputados, que deve agora discutir e deliberar sobre a proposta. A iniciativa foi encaminhada à Comissão de Ciência e Tecnologia, Comunicação e Informática (CCTCI) e está sob os auspícios do

Grupo de Trabalho de Aperfeiçoamento da Legislação Brasileira na Internet (GTNET) a quem cabe analisar e elaborar parecer sobre o PL 2.630/2020 e as 79 propostas a ele apensadas até o momento. Na casa revisora, o texto está sob a relatoria do deputado Orlando Silva (PCdoB-SP) e foi objeto de ampla discussão em dois ciclos de audiências públicas, ouvindo representantes da academia, da sociedade civil e do setor privado (BRASIL, 2021d).

3.2 DISPOSIÇÕES PRELIMINARES

O *caput* do art. 1º sintetiza o conteúdo da proposta e delimita seu escopo, indicando, nos seguintes termos, que a referida lei:

“[...] estabelece normas relativas à transparência de redes sociais e de serviços de mensagens privadas, sobretudo no tocante à **responsabilidade dos provedores pelo combate à desinformação e pelo aumento da transparência na internet**, à transparência em relação a conteúdos patrocinados e à atuação do poder público, bem como estabelece sanções para o descumprimento da lei. (BRASIL, 2020b, p. 1, grifo do autor)”

O parágrafo primeiro do mesmo dispositivo estabelece que as regras ali previstas não se aplicarão aos provedores de redes sociais e de serviços de mensageria privada que possuam “menos de 2.000.000 (dois milhões) de usuários registrados” no país (BRASIL, 2020b, p. 1). Desse modo, apenas serviços com uma quantidade superior de membros estarão obrigados cumprir as novas regras, caso este texto entre em vigência.

Logo em seguida, foram elencados nove princípios que alicerçam as disposições da referida proposta legislativa:

“Art. 3º Esta Lei será pautada pelos seguintes princípios:
I – **liberdade de expressão e de imprensa;**
II – **garantia dos direitos de personalidade, da dignidade, da honra e da privacidade do indivíduo;**
III – **respeito ao usuário em sua livre formação de preferências políticas e de uma visão de mundo pessoal;**
IV – **responsabilidade compartilhada pela preservação de uma esfera pública livre, plural, diversa e democrática;**

V – garantia da confiabilidade e da integridade dos sistemas informacionais;

VI – promoção do acesso ao conhecimento na condução dos assuntos de interesse público;

VII – acesso amplo e universal aos meios de comunicação e à informação;

VIII – proteção dos consumidores; e

IX – transparência nas regras para veiculação de anúncios e conteúdos pagos. (BRASIL, 2020b, p. 1-2, grifo do autor)”

Posteriormente, o legislador enuncia cinco objetivos a serem alcançados por meio desta proposta. Destaca-se, neste particular, os incisos II e III que elencam como alvos “o impedimento da censura no ambiente online” e “a garantia do contraditório e da ampla defesa”, a saber:

“Art. 4º Esta Lei tem como objetivos:

I – o fortalecimento do processo democrático por meio do combate ao comportamento inautêntico e às redes de distribuição artificial de conteúdo e do fomento ao acesso à diversidade de informações na internet no Brasil;

II – a defesa da liberdade de expressão e o impedimento da censura no ambiente online;

III – a busca por maior transparência das práticas de moderação de conteúdos postados por terceiros em redes sociais, com a garantia do contraditório e da ampla defesa; e

IV – a adoção de mecanismos e ferramentas de informação sobre conteúdos impulsionados e publicitários disponibilizados para o usuário. (BRASIL, 2020b, p. 2, grifo do autor)”

Convém ainda destacar algumas das definições previstas no art. 5º, pertinentes à abordagem regulatória procedimental da governança de conteúdo na internet, quais sejam:

“Art. 5º Para os efeitos desta Lei, considera-se:

[...]

V – conteúdo: dados ou informações, processados ou não, contidos em qualquer meio, suporte ou formato, compartilhados em redes sociais ou serviços de mensageria privada, independentemente da forma de distribuição, publicação ou transmissão utilizada pela internet;

[...]

VIII – rede social: aplicação de internet que se destina a realizar a conexão de usuários entre si, permitindo e tendo como centro da atividade a comunicação, o compartilhamento e a disseminação de conteúdo em um mesmo sistema de informação, através de contas conectadas ou acessíveis entre si de forma articulada; e

[...]

IX – serviço de mensageria privada: aplicação de internet que viabiliza o envio de mensagens para destinatários certos e determinados, inclusive protegidas por criptografia de ponta a ponta, a fim de que somente remetente e destinatário da mensagem tenham acesso ao seu conteúdo, excluídas aquelas prioritariamente destinadas a uso corporativo e os serviços de correio eletrônico. (BRASIL, 2020b, p. 2-3)”

O parágrafo único do referido artigo ainda afasta “empresas jornalísticas” do escopo da lei, caso estas viessem a se enquadrar na definição de “provedores de redes sociais”.

Por fim, é digna de nota a decisão do relator de não incluir uma definição específica sobre “desinformação”, ainda que o combate a essa prática seja um dos principais objetivos declarados.

No relatório, o autor do substitutivo justificou a decisão por considerar que há um linha tênue entre manifestações com proteção constitucional e prática de desinformação. Para o senador, “a atividade de rotular determinada informação como falsa é, em si, uma manifestação de opinião” (BRASIL, 2020c, p. 6).

3.3 OS PROCEDIMENTOS DE MODERAÇÃO

O art. 12 do PL 2.630/2020 prevê novos direitos e garantias processuais aos usuários das redes sociais na moderação de conteúdo. Não é redundante sublinhar que tais disposições constituem também novas obrigações aos provedores de aplicação de internet incluídos na proposta (BRASIL, 2020b, p. 5-6).

3.3.1 Direito à informação e à mecanismos de recurso e devido processo

O *caput* do décimo segundo artigo estabelece o direito à informação e o direito à liberdade de expressão em duas situações distintas na governança de conteúdo na internet: o processo de *elaboração* dos termos de uso das plataformas, e a *aplicação* dessas políticas (BRASIL, 2020b, p. 5).

De acordo com essa acepção, os usuários passarão a ter a prerrogativa não apenas de serem informados a respeito de como as normas e padrões da comunidade são definidas e aplicadas, mas terão direito a voz neste particular.

O dispositivo ainda prevê, de forma genérica, a disponibilização de mecanismos de recurso e devido processo por parte dos provedores de aplicação. Leia-se:

“Art. 12. Os provedores de aplicação de internet submetidos a esta Lei devem garantir o direito de acesso à informação e à liberdade de expressão de seus usuários nos processos de elaboração e aplicação de seus termos de uso, **disponibilizando mecanismos de recurso e devido processo**. (BRASIL, 2020b, p. 5, grifo do autor)”

3.3.2 Direito à notificação, à decisão fundamentada e ao contraditório

O § 1º do artigo 12 prescreve que a imposição de medidas de moderação, tomadas com base nas regras da própria plataforma ou na lei, deve ser necessariamente acompanhada da devida notificação ao(s) usuário(s) afetado(s). Isso também deve acontecer “em caso de denúncia” contra uma publicação ou usuário. A mensagem da notificação, por sua vez, deve conter as razões pelas quais a decisão foi tomada (ou no caso da denúncia, a acusação feita), além de informar prazos pertinentes e procedimentos de contestação (BRASIL, 2020b, p. 5).

Observa-se, portanto que esse dispositivo introduz o direito à notificação dos usuários, o direito à fundamentação das decisões, e o direito ao contraditório por garantir a oportunidade de contestação às partes afetadas (BRASIL, 2020b, p. 5).

O texto ainda enuncia um rol de hipóteses nas quais haveria uma dispensa à notificação aos usuários. Incluem-se na lista situações em que fosse *verificado* risco de “dano imediato de difícil reparação”, ameaça à “segurança da informação ou do usuário”, perigo “de violação a direitos de crianças e adolescentes”, além da possibilidade “de grave comprometimento da usabilidade, integridade ou estabilidade da aplicação”. O dispositivo ainda desobriga a notificação caso haja risco do cometimento de crimes de racismo, previstos na Lei nº 7.716 (BRASIL, 2020b, p. 5-6). Leia-se:

“§ 1º Em caso de denúncia ou de medida aplicada em função dos termos de uso das aplicações ou da presente Lei que recaia sobre conteúdos e contas em operação, **o usuário deve ser notificado sobre a fundamentação, o processo de análise e a aplicação da medida, assim como sobre os prazos e procedimentos para sua contestação.**

§ 2º Os provedores dispensarão a notificação aos usuários se verificarem risco:

- I – de dano imediato de difícil reparação;
- II – para a segurança da informação ou do usuário;
- III – de violação a direitos de crianças e adolescentes;
- IV – de crimes tipificados na Lei nº 7.716, de 5 de janeiro de 1989;
- V – de grave comprometimento da usabilidade, integridade ou estabilidade da aplicação. (BRASIL, 2020b, p. 5-6, grifo do autor)”

3.3.3 Direito ao recurso

Tanto o *caput* quanto o parágrafo terceiro do dispositivo em apreço enunciam de forma explícita a garantia ao acesso de mecanismos de recurso. Desse modo, passará a existir o direito de recorrer de decisões de bloqueio de conta, ou mesmo da remoção de um determinado conteúdo, nos seguintes termos:

“§ 3º Deve ser garantido pelo provedor o direito de o usuário recorrer da indisponibilização de conteúdos e contas. (BRASIL, 2020b, p. 6)”

3.3.4 Direito à reparação por danos decorrentes de decisão equivocada

Outrossim, na hipótese de um juízo equivocado por parte da plataforma, no qual se tenha considerado que um determinado conteúdo transgrediu os padrões de conduta ou a própria lei, e, dessa decisão decorra algum dano, caberá ao provedor de redes sociais reestabelecer a justiça e reparar o mal causado, “no âmbito e nos limites técnicos do serviço”. Não se trata, portanto, de responsabilização civil, penal ou administrativa. Observe-se:

“§ 4º Havendo dano decorrente da caracterização equivocada de conteúdos como violadores dos padrões de uso de aplicações ou do disposto na presente Lei, caberá ao provedor de redes sociais repará-lo, no âmbito e nos limites técnicos do serviço. (BRASIL, 2020b, p. 6)”

3.3.5 Direito a prazo de defesa maior no caso de *deepfakes*

O parágrafo quinto, por sua vez, disciplina um prazo distinto para a apresentação de defesa em relação a conteúdos que “utilizem imagem ou voz manipuladas para imitar a realidade”, as chamadas *deepfakes*.

A períclope ainda estabelece que, para a incidência do benefício, é necessário que o conteúdo tenha sido veiculado com o objetivo de “induzir a erro acerca da identidade de candidato a cargo público”. Trata-se, portanto, de uma regra com repercussões de natureza eleitoral.

O excerto ainda ressalva, contudo, que o prazo estendido não se aplica à conteúdos que caracterizem humor ou paródia. Leia-se:

“§ 5º O prazo de defesa será diferido nos casos de conteúdo que use imagem ou voz manipuladas para imitar a realidade, com o objetivo de induzir a erro acerca da identidade de candidato a cargo público, ressalvados o ânimo humorístico ou de paródia. (BRASIL, 2020b, p. 6)”

3.3.6 Direito de resposta proporcional ao agravo

O último parágrafo do art. 12 estabelece que, nas situações em que determinado conteúdo ofensivo a outrem seja considerado indevido pelo provedor de rede social, deve-se garantir ao ofendido o direito de resposta proporcional “na medida e alcance” do conteúdo que o injuriou.

“§ 6º A decisão do procedimento de moderação deverá assegurar ao ofendido o direito de resposta na mesma medida e alcance do conteúdo considerado inadequado. (BRASIL, 2020, p. 6)”

3.4O CONSELHO DE TRANSPARÊNCIA E REPONSABILIDADE NA INTERNET (CTRI)

O Capítulo IV do PL 2.630/2020 prevê a criação de um conselho de caráter consultivo denominado Conselho de Transparência e Responsabilidade na Internet (CTRI). O referido órgão, de acordo com o art. 25, deverá ser instituído pelo Congresso Nacional “em até 60 (sessenta) dias contados da publicação da lei”, e terá como atribuições “a realização de estudos, pareceres e recomendações sobre liberdade, responsabilidade e transparência na internet.” (BRASIL, 2020, p. 9-10).

A esse órgão também caberá, de acordo com o art. 25, as responsabilidades de:

“[...] II – **elaborar código de conduta** para redes sociais e serviços de mensageria privada, a ser avaliado e aprovado pelo Congresso Nacional, aplicável para a garantia dos princípios e objetivos estabelecidos nos arts. 3º e 4º desta Lei, **dispondo sobre fenômenos relevantes no uso de plataformas por terceiros, incluindo, no mínimo, desinformação, discurso de incitação à violência, ataques à honra e intimidação vexatória;**

III – avaliar os dados constantes nos relatórios de que trata o art. 13 desta Lei;

IV – publicar indicadores sobre o cumprimento dos códigos de conduta pelo setor;

V – **avaliar a adequação das políticas de uso adotadas pelos provedores de redes sociais e de serviços de mensageria privada;**

VI – organizar, anualmente, conferência nacional sobre liberdade, responsabilidade e transparência na internet;

[...]

VIII – avaliar os procedimentos de moderação adotados pelos provedores de redes sociais, bem como sugerir diretrizes para sua implementação;

IX – promover estudos e debates para aprofundar o entendimento sobre desinformação, e o seu combate, no contexto da internet e das redes sociais;

X – certificar a entidade de autorregulação que atenda aos requisitos previstos nesta Lei; e

XI – estabelecer diretrizes e fornecer subsídios para a autorregulação e para as políticas de uso dos provedores de redes sociais e de serviços de mensageria privada. (BRASIL, 2020b, p. 10-11, grifo do autor).”

O Conselho será composto por “21 (vinte e um) conselheiros, com mandato de 2 (dois) anos, admitida 1 (uma) recondução”. Deverão fazer parte do órgão representantes do Senado Federal (1), da Câmara dos Deputados (1), da sociedade civil (5), da academia e comunidade técnica (2), dos provedores de acesso, aplicações e conteúdo da internet (2), do setor de comunicação social (2), do setor de

telecomunicações (1), do CNJ (1), do CNMP (1), do CNCPC (1), do CGI (1), do Departamento da Polícia Federal (1), da Anatel (1), e do Conar (1) (BRASIL, 2020b, p. 10-11).

Para além de outras disposições burocráticas, o texto vincula as despesas com a instalação e o funcionamento do CTRI “à conta do orçamento do Senado Federal” (BRASIL, 2020b, p. 10).

3.5 UMA INSTITUIÇÃO DE AUTORREGULAÇÃO REGULADA

A proposta legislativa ainda *recomenda*, em seu art. 30, a criação de uma “instituição de autorregulação voltada à transparência e à responsabilidade no uso da internet” (BRASIL, 2020b, p. 11).

O ente, a ser eventualmente constituído pelos próprios provedores de redes sociais e de serviços de mensageria privada, teria como atribuições:

“I – criar e administrar plataforma digital voltada à transparência e à responsabilidade no uso da internet, que contenha **regras e procedimentos para decidir sobre a adoção de medida informativa**, atendendo ao disposto nesta Lei;

II – assegurar a independência e a especialidade de seus analistas;

III – disponibilizar serviço eficiente de atendimento e encaminhamento de reclamações;

IV – estabelecer requisitos claros, objetivos e acessíveis para a participação dos provedores de redes sociais e serviços de mensageria privada;

V – incluir em seu quadro uma ouvidoria independente com a finalidade de receber críticas e avaliar as atividades da instituição; e

VI – desenvolver, em articulação com as empresas de telefonia móvel, boas práticas para suspensão das contas de usuários cuja autenticidade for questionada ou cuja inautenticidade for estabelecida. (BRASIL, 2020, p. 11-12, grifo do autor).”

Como um mecanismo de autorregulação, a referida instituição poderá aprovar “resoluções e súmulas de modo a regular seus procedimentos de análise”, conforme prevê o § 3º.

De acordo com o parecer, a escolha desse modelo institucional foi feita para evitar a criação de “limitações indevidas a outros tipos de aplicações de internet e com

isso causar insegurança jurídica e impactos econômicos indesejados” (BRASIL, 2020c, p. 9).

3.6 AS SANÇÕES

Caso descumpram as novas regras, a autoridade judicial poderá, “sem prejuízo das demais sanções civis, criminais ou administrativas”, sujeitar os provedores de redes sociais e de serviços de mensageria privada a:

- I – advertência, com indicação de prazo para adoção de medidas corretivas; ou
- II – multa de até 10% (dez por cento) do faturamento do grupo econômico no Brasil no seu último exercício. (BRASIL, 2020b, p. 12).”

3.7 DISPOSIÇÕES FINAIS RELEVANTES

Por fim, o art. 35 do PL 2.630/2020 define que a lei entrará em vigor após 90 dias de sua publicação oficial, com exceção dos artigos 25 a 29, correspondes ao Conselho de Transparência e Responsabilidade na Internet, que terão vigência imediata com a veiculação da norma.

4 ANÁLISE DO PL 2.630/2020 À LUZ DA PERSPECTIVA PROCEDIMENTAL DE REGULAÇÃO DAS GRANDES PLATAFORMAS DIGITAIS

O objetivo deste capítulo é eminentemente exploratório. Com ele, buscar-se-á analisar as disposições do PL 2.630/2020 delimitadas no capítulo anterior, tendo como prisma de análise a perspectiva regulatória baseada no devido processo, conforme exposta no primeiro capítulo. Também são utilizados como parâmetro os padrões internacionais para a proteção da liberdade de expressão no ambiente virtual.

Inicialmente, proceder-se-á a uma análise panorâmica do PL 2.630/20, o processo legislativo e seus principais aspectos. Em seguida, serão objeto de consideração os procedimentos de moderação dispostos na proposta legislativa supracitada. Ademais, analisar-se-á as inovações estruturais previstas no texto aprovado pelo Senado Federal. Por fim, são feitas recomendações ao Congresso Nacional relativas à iniciativa legiferante em apreço.

4.1 CONSIDERAÇÕES INICIAIS

Por disciplinar um tema de elevada complexidade técnica e alta sensibilidade jurídico-constitucional, um marco regulatório que busque zelar por um ambiente virtual livre, responsável e transparente, não deve tramitar de forma açodada e sem um amplo debate público sobre suas disposições.

Infelizmente, não foi o que aconteceu com o texto do PL 2.630/20 aprovado pelo Senado Federal. Da apresentação da versão inicial pelo senador Alessandro Vieira (Cidadania-SE) em 13/05/20, até a aprovação do substitutivo pelo Plenário da Câmara Alta em 30/06/20, decorreram apenas sete semanas.

Ao longo deste período não foram convocadas audiências públicas para discutir a proposta, e o texto não foi enviado a comissões temáticas para análise e discussão. Em um período curto, foram apresentados quatro relatórios distintos, fortemente criticados por acadêmicos e ativistas³⁷. Após intensa pressão da sociedade

³⁷ Uma das versões alternativas apresentadas pelo relator chegou a ser considerada como uma das “piores de leis de internet do mundo” por defensores da liberdade de expressão nas redes (PALLERO *et al*, 2020b; CDR, 2020).

civil, foi aprovada uma proposta menos problemática que seguiu ao escrutínio da casa revisora.

Distintamente, a Câmara dos Deputados promoveu uma ampla discussão sobre o Projeto de Lei nº 2630/2020 em dois ciclos distintos de audiências públicas. Sob os auspícios do Grupo de Trabalho de Aperfeiçoamento da Legislação Brasileira na Internet (GTNET) e da Comissão de Ciência e Tecnologia, Comunicação e Informática (CCTCI), foram ouvidos representantes dos mais diversos setores da sociedade a respeito do tema.

Espera-se que, em breve, os parlamentares possam discutir, aperfeiçoar e aprovar um marco regulatório voltado às grandes plataformas de conteúdo que seja robusto e efetivamente proteja direitos fundamentais, garantindo um ambiente digital livre, responsável, aberto e seguro.

4.2 DISPOSIÇÕES PRELIMINARES

4.2.1 Escopo

Em seus prolegômenos, o texto legislativo limita a aplicação da norma a “provedores de redes sociais e de serviços de mensageria privada” com mais de dois milhões de usuários nacionais registrados. Percebe-se que o legislador brasileiro utilizou o mesmo critério adotado pela NetzDG (ALEMANHA, 2017). Trata-se de um referencial que demanda uma análise factual mais criteriosa, visto que há uma enorme diferença em termos de alcance e onerosidade para as plataformas de grande porte (ex.: Facebook, Instagram, Twitter, YouTube, TikTok, etc.) e outras de menor porte, como o Reddit, que já conta com somente 7,51 milhões de usuários nacionais (STATISA, 2021).

Sob esse prisma, recomenda-se a adoção de um critério progressivo para regular as plataformas digitais. Isso significa que quanto maior seja o tamanho, o alcance e a posição de dominância de um intermediário, maior deverá ser o rigor regulatório adotado. Isso poderá ajudar a evitar a imposição de onerosidades excessivas em *players* de menor relevância. Ademais, as plataformas que moderam

conteúdo de bilhões de usuários, impactam de forma mais significativa os direitos fundamentais na internet. Os *gatekeepers* da internet detêm um enorme controle sobre os canais de acesso à informação, comunicação e deliberação no ambiente virtual. Esse poder oligopolístico deve ser balanceado com mecanismos de *accountability*, transparência, devido processo, e fomento à competição.

Conquanto seja fundamental delimitar o escopo de incidência, observa-se que o legislador optou por não incluir certos intermediários como mecanismos de busca (ex.: Google etc.), que constituem importantes fontes de acesso à informação, opiniões, e bens culturais. Mais do que isso, os buscadores tomam decisões substanciais e ubíquas sobre conteúdo, impactando diretamente o debate público e o exercício da liberdade de expressão (PALLERO *et al*, 2020, p. 7). Pode ser interessante considerar sua inclusão.

Por outro lado, o texto inclui em seu escopo serviços de mensageria privada, que, conquanto possam ser utilizados para a comunicação em massa, exercem como principal função a comunicação interpessoal e não costumam fazer moderação de conteúdo, o que limita as possibilidades de regulação (PALLERO *et al*, 2020, p. 7). Deve-se, portanto, avaliar a adequação dessa categoria no escopo da lei.

A esse respeito, seria relevante incluir todos os provedores de serviços online que atuem como “intermediários ou plataformas de armazenamento, busca ou troca de informações, opiniões, expressões e outros conteúdos gerados por seus usuários e que realizam algum tipo de curadoria ou moderação desses conteúdos” (PALLERO *et al*, p. 7). São precisamente estes sujeitos do processo comunicativo no ambiente virtual que exercem a relevante função de controle de conteúdo que impacta diretamente o exercício da liberdade de expressão. É a estes atores, em especial, que compete deveres de ética, responsabilidade e confiabilidade.

4.2.2 Objetivos

Outrossim, o PL 2.630/2020 designa como objetivos “a *defesa da liberdade de expressão* e o impedimento da “*censura no ambiente online*” (art. 4, II) além da “*busca por maior transparência das práticas de moderação de conteúdo [...]*, com a garantia do *contraditório e da ampla defesa*” (art. 4, III, grifo do autor). O projeto também declara perseguir o “*fortalecimento do processo democrático por meio do combate ao*

comportamento inautêntico [...] e do fomento ao acesso à *diversidade de informações* na internet no Brasil” (art. 4, I, grifo do autor) bem como a “adoção de mecanismos e ferramentas de *informação sobre conteúdos impulsionados e publicitários* [...]” (art. 4, IV, grifo do autor).

Balkin (2018) argumenta que qualquer proposta regulatória voltada para a proteção da liberdade de expressão na era digital deve se direcionar a dois objetivos centrais: (i) a prevenção e a limitação, no limite do possível, da prática da censura colateral e da censura prévia; e (ii) a proteção dos usuários contra a vigilância a manipulação.

Para atingir tais objetivos, o autor propõe, entre outras coisas, o estabelecimento de obrigações típicas do *devido processo* à moderação de conteúdo nas plataformas (ex.: direito à recurso, direito a decisões fundamentadas, direito à notificação etc.) e o reconhecimento dos deveres de confidencialidade, de cuidado e de lealdade para com os usuários que decorrem do papel de “fiduciários da informação” que deve ser exercido pelas plataformas (BALKIN, 2018).

Bowers e Zittrain (2020, p. 8), por sua vez, defendem que o principal objetivo dos novos modelos regulatórios deve ser a construção de mecanismos consensuais e legítimos para a tomada de decisão em questões difíceis da governança de conteúdo. Ou seja, ao invés de focar propriamente em dar as respostas para as questões difíceis (juízo substancial), reguladores e legisladores devem se concentrar na construção de um *procedimento* legítimo e amplamente reconhecido para a tomada dessas decisões.

Para isso, os autores sugerem que sejam instituídas obrigações aos intermediários que decorram do seu reconhecimento como “fiduciários do conteúdo” gerado por usuários (ex.: obrigações do devido processo na moderação de conteúdo). Também sugerem que decisões sensíveis, capazes de estabelecer precedentes, sejam delegadas a entes externos e independentes, com o distanciamento necessário (BOWERS & ZITTRAIN, 2020).

Face a tais considerações, é de se observar, inicialmente, que o PL 2.630/2020 incorpora em parte as preocupações dos autores de referência. Isso porque define como alvo o combate à “censura no ambiente online” e a garantia da “ampla defesa e o contraditório” na moderação de conteúdo. O texto também acrescenta às suas finalidades a adoção de práticas que podem reduzir a manipulação e a vulnerabilidade

dos usuários como mecanismos de transparência, o “combate ao comportamento inautêntico” e a promoção da “diversidade de informações” na internet.

É preciso, porém, que os mecanismos regulatórios desenhados pelo legislador sejam efetivamente capazes de atingir tais objetivos. O esforço de engenharia regulatória deve compatibilizar mecanismos de transparência, *accountability*, devido processo, direito à defesa, não discriminação, necessidade e proporcionalidade com as garantias constitucionais à liberdade de expressão.

4.3 AS REGRAS DE MODERAÇÃO

O décimo segundo artigo do PL 2.630/20 prevê a criação de novos direitos e garantias de natureza procedimental aos usuários na moderação de conteúdo. Dentre as inovações legislativas estão o direito à informação, o direito ao devido processo, o direito à notificação, o direito à decisão fundamentada, o direito ao contraditório, o direito ao recurso e o direito de resposta.

4.3.1 Transparência e o direito à informação

O *caput* do art. 12 do PL 2.630/20 inova positivamente ao prever de forma explícita o direito à informação tanto nos “processos de elaboração” quanto na “aplicação” dos termos de uso dos provedores de aplicação de internet. Tal provisão é bem-vinda e ecoa as recomendações para uma regulação democrática de plataformas digitais.

O direito à informação no contexto das plataformas digitais não se satisfaz com a mera disponibilização das regras e diretrizes de funcionamento dos espaços virtuais. Esta garantia configura, sobretudo, uma obrigação de clareza, acessibilidade, precisão e compreensibilidade a ser cumprida pelas empresas ao dispor sobre suas normas.

Além de publicar suas políticas em formatos acessíveis e em lugares de fácil alcance aos usuários, deve-se empregar linguagem clara e inteligível, direcionadas ao público a quem se destinam. Também é mister que as companhias mantenham tais provisões sempre atualizadas, à medida que evoluem e sofrem alterações.

Para além disso, mais do que expor as regras em si, o direito à informação alcança o próprio processo de criação e implementação das normas privadas. Trata-se, nesta acepção, de um importante passo para aperfeiçoar os termos de serviços, que são, em grande medida, elaborados e aplicados sem transparência e de forma soberana pelas empresas.

Ao lançar luz sobre o processo de criação e aplicação das “leis virtuais”, abre-se caminho para contrabalancear o poder ilimitado e discricionário exercido pelas grandes plataformas no julgamento de conteúdos em suas comunidades virtuais. Essa disposição será importante, sobretudo, para trazer mais clareza nas hipóteses em que os termos de serviços são vagos, ou excessivamente amplos, podendo ser interpretados de forma a limitar indevidamente o exercício da liberdade de expressão dos usuários.

4.3.2 O direito ao devido processo e ao recurso

A proposta legislativa em apreço elege como um de seus objetivos fundamentais “a garantia do contraditório e da ampla defesa” (art. 4, III) e o *caput* do art. 12 prevê de forma genérica a disponibilização de “mecanismos de recurso e devido processo”. Mais adiante, o mesmo dispositivo prescreve o direito de recorrer da “indisponibilização de conteúdos e contas”, sem, mais uma vez, elencar balizas mínimas para a materialização dessa garantia.

Conquanto seja positiva a indicação dessas garantias, a não especificação de parâmetros elementares, pode abrir espaço para procedimentos disfuncionais que não sejam capazes de assegurar efetivamente guaridas processuais aos usuários. Um exemplo disso é a não previsão de prazos mínimos de defesa, o que não dá efetividade procedimental à norma.

Além disso, o art. 12 parece alheio ao modelo de responsabilização civil de intermediários disciplinado pelo Marco Civil da Internet³⁸, que trata precisamente sobre a exclusão de conteúdo e/ou contas da internet. Seria interessante, nesse sentido, que o legislador tomasse como referência o eixo previsto no MCI, prevendo expressamente a hipótese de remoção para o cumprimento de ordem judicial.

³⁸ Artigos 19 e 21 da Lei nº 12.965, de 23 de abril de 2014.

4.3.3 O direito à notificação e à decisão fundamentada

O parágrafo primeiro do art. 12 contempla um desenho regulatório que inclui o direito à notificação em caso de denúncia ou da aplicação de alguma medida, bem como o direito à fundamentação das decisões de moderação que lhe digam respeito, além de prescrever procedimentos de contestação típicos do direito ao contraditório. Tais mecanismos regulatórios constituem importante degrau para coibir abusos e arbitrariedades no exercício da liberdade de expressão na internet.

Desse modo, como regra geral, não deveria ser permitida a derrubada de publicações ou outros conteúdos gerados por usuários sem antes notificá-los a respeito da medida, fornecendo as devidas justificativas, e a possibilidade de apelação. Nessa dinâmica, mesmo diante de uma decisão arbitrária, os usuários terão assegurado o direito de defesa e a oportunidade de questionar abusos.

Um ponto problemático, contudo, diz respeito ao § 2º do art. 12, que prevê um rol de cinco situações em que seria possível remover conteúdo sem a necessidade de notificar os usuários afetados. É possível que tal disposição, na forma como foi redigida, constitua uma espécie de obrigação de remoção imediata de conteúdo nos casos especificados, capaz agravar ainda mais situações de censura prévia e colateral.

Mais do que isso, a textura aberta de alguns dos enunciados (I, II, V) do segundo parágrafo pode dar azo a interpretações excessivamente subjetivas e abrangentes das situações que se encaixariam naquele contexto. Isso pode aumentar a discricionariedade das plataformas, dando a elas uma espécie de “sinal verde” para a ‘tirar de circulação’ uma ampla variedade de conteúdo, sem que as devidas explicações sejam apresentadas (AL SUR, 2021, p. 14).

Na presente redação, este dispositivo pode ensejar ainda mais arbitrariedades, colocando em especial risco a liberdade de expressão de grupos e indivíduos vulnerabilizados. Se o modo pouco transparente com que as plataformas aplicam suas políticas de moderação de conteúdo já permite a ocorrência desse tipo de situação, quanto mais se houver uma salvaguarda legal que dê resguardo a arbitrariedades (AL SUR, 2021, p. 14).

4.3.4 Outros casos: reparação de danos, direito a resposta e *deepfakes*

É também salutar a previsão de um direito à reparação por danos decorrentes de decisão equivocada tomada pela plataforma na moderação de conteúdo. A norma ainda prevê que essa reparação deverá ocorrer “nos âmbito e nos limites técnicos do serviço”. Ou seja, a empresa deve lançar mão de todos os esforços necessários e ao alcance de suas capacidades técnicas para reparar os prejuízos causados pelo equívoco.

Não apenas isso, a proposta legislativa inova ao prever uma hipótese processual específica para conteúdos que “utilizem imagem ou voz manipuladas para imitar a realidade”. São as *deepfakes*, montagens de vídeo ou áudio que reproduzem com perfeição o rosto ou a voz de uma determinada pessoa, sendo, em muitos casos, praticamente impossível a diferenciação da pessoa real. Esse recurso se popularizou nos últimos anos sobretudo para fins humorísticos.

No entanto, caso o *deepfake* de um “candidato a cargo público” seja veiculado nas redes sociais com o objetivo de induzir o público a erro em relação a identidade postulante, o §5º prevê que as plataformas deverão, ao adotar medidas de moderação, providenciar um prazo de defesa diferenciado aos implicados. Esse dispositivo pode garantir o tempo necessário para que os indivíduos afetados demonstrem a falsificação, preservando-lhes o direito à imagem e à honra.

Para além disso, seria interessante que o legislador considerasse outras hipóteses que demandam prazo diferenciado, seja para proporcionar uma defesa rápida, como em casos de incitação à violência, ou outras hipóteses de prazo alongado.

4.4 AS INOVAÇÕES INSTITUCIONAIS

4.4.1 O Conselho de Transparência e Responsabilidade na Internet (CTRI)

O PL 2.630/2020 dedicou um de seus capítulos para disciplinar a criação e o funcionamento de um conselho consultivo que tem como atribuições, entre outras

coisas, a “elaboração de um código de conduta” para os provedores de redes sociais e aplicações de mensageria privada (art. 25, II). O novo conselho também deverá ser responsável por analisar a “adequação das políticas de uso” adotadas pelas plataformas às suas próprias recomendações, além de “avaliar os procedimentos de moderação” e “sugerir diretrizes para sua implementação” (art. 25, V e VIII).

É muito bem-vinda a criação de um conselho multissetorial capaz de se manifestar e avaliar em questões relevantes que digam respeito à governança de conteúdo na internet. É também louvável a presença de representantes dos mais diversos setores interessados no órgão colegiado, incluindo membros da sociedade civil. No entanto, gera estranheza a previsão de dois representantes de órgãos policiais (um representante do Conselho Nacional dos Chefes de Polícia Civil e um do Departamento da Polícia Federal), o que parece indicar uma compreensão equivocada sobre o significado da moderação de conteúdo, e sua relação com uma postura policial. A presença dessa categoria, que merece o devido respeito institucional, parece incompatível com os propósitos do CTRI.

Além disso, o conselho constituirá um organismo sem poder normativo, ou seja, de caráter eminentemente recomendatório. Caso as empresas de internet optem por não seguir o código de conduta criado, ou por não obedecer às recomendações do ente em questão, não haverá meios de garantir o *enforcement* das medidas. Por mais que o colegiado possa emitir juízos sobre os procedimentos de moderação, este não será capaz de assegurar na prática a proteção do direito à liberdade de expressão, nem tampouco a adoção das medidas adequadas ao resguardo do interesse público em questões nas quais isso seja necessário.

Uma forma possível de fortalecer o conselho seria estabelecer uma nova obrigação para que os provedores de redes sociais apresentem relatórios de transparência, nos quais se informe especificamente sobre as normas de moderação de conteúdo e estatísticas sobre como estas têm sido aplicadas. O órgão colegiado poderia ser responsável por definir os parâmetros e especificidades a serem dispostos nesse documento.

Além disso, conquanto a proposta legislativa preveja a origem dos recursos a serem destinados para a criação e o funcionamento do CTRI, o art. 29 vincula as despesas do Conselho ao orçamento do Senado Federal. Não se trata, portanto, de um órgão com independência orçamentária. Essa “engenharia institucional” levanta

questionamentos a respeito de sua independência e capacidade de tomar decisões com o devido distanciamento. Por não ter orçamento próprio, o Conselho pode ficar sujeito a pressões políticas indevidas.

Por fim, embora a iniciativa legiferante elenque o “combate à desinformação” como um de seus alvos principais (art. 1), o legislador esquivou-se da tarefa de definir o termo, delegando essa responsabilidade ao futuro órgão colegiado, que, sem poderes normativos, deverá redigir um código de conduta endereçando tais práticas (art. 25, II). Não há garantias, nessa construção, de que esses males serão efetivamente combatidos (AL SUR, 2021, p. 15).

4.4.2 Uma instituição de autorregulação regulada

Há ainda uma provisão peculiar na proposta legislativa. O texto aprovado pelo Senado recomenda aos provedores de redes sociais e de serviços de mensageria privada que criem, de iniciativa própria, uma “instituição de autorregulação voltada à transparência e à responsabilidade no uso da internet” (art. 30). Esse ente consagraria a opção por um modelo de “autorregulação regulada”³⁹, da governança privada de conteúdo online no Brasil.

De acordo com o art. 30, a instituição teria o condão de decidir sobre a adoção de “medidas informativas” de seus membros. Outra atribuição do organismo seria assegurar a independência de seus especialistas e analistas.

Por um lado, a adoção de uma postura menos intervencionista é positiva pois evita a criação de estímulos para a censura prévia e colateral. Obrigações regulatórias que forcem as empresas a monitorar ostensivamente o conteúdo de seus usuários a fim de identificar supostas violações à lei, ou impedir que elas aconteçam geram disfuncionalidades, conforme apontou Balkin (2018).

Por outro lado, um modelo regulatório que não estabeleça freios e contrapesos à governança privada, deixando os usuários à mercê de interesses econômicos das grandes corporações, coloca em risco direitos individuais e o bem-estar coletivo, conforme observado a seguir:

³⁹ No modelo de autorregulação regulada, “algumas premissas são dadas pelo Estado, cabendo às empresas buscar códigos de conduta internos para melhor se adequarem à nova realidade” (AGAPITO *et al*, 2017, p. 368; *apud* SCAFF *et al*, 2014)

“[...] o modelo de autorregulação que tem prevalecido até agora apresenta riscos similares ao efetivo exercício de direitos humanos básicos. **Algumas poucas corporações centralizaram e concentraram o poder de administrar a circulação, troca ou busca de informações e opiniões e exercem esse poder de forma arbitrária e sem nenhum mecanismo de *accountability* às instituições que garantem direitos. Esse cenário representa um risco ao efetivo exercício de direitos, um risco que vai crescendo à medida que tal punhado de empresas reforça sua posição dominante no mercado e desenvolve tecnologias opacas para a governança da informação.** (PALLERO *et al*, 2020, p. 7, grifo do autor)”

A tomada de decisões sensíveis e paradigmáticas ao exercício da liberdade de expressão na internet, que interessam diretamente à coletividade não deve estar sujeita às forças e pressões de interesses políticos ou privados. Esse papel deve caber a uma organismo idôneo, com garantias reais de autonomia técnica e decisória, independência orçamentária e imparcialidade. Mais do que isso, esse ente, caso seja criado, deve ter condições de assegurar o cumprimento de suas decisões.

Recentemente novas soluções autorregulatórias têm surgido, como o Facebook Oversight Board, que prevê, em tese, prerrogativas de independência financeira, autonomia técnica e decisões vinculantes. E conquanto as decisões e recomendações recentes do organismo tenham se mostrado positivas, é preciso mais tempo para avaliar seu impacto real na prática e dinâmica da governança privada de conteúdo (ARCHEGAS *et al*, 2021, p. 24-26).

4.5 SANÇÕES

Caso não cumpram as regras dispostas no PL 2.630/2020, os provedores de redes sociais e as aplicações de mensageria privada estarão sujeitas a sanções como advertências e multas, a serem aplicadas por autoridades judiciais.

Porém, tais instrumentos de coerção são ainda limitados para endereçar as peculiaridades que caracterizam os problemas da governança privada de conteúdo digital. Outras medidas de reparação seriam salutares, como a promoção de iniciativas voltadas ao fomento da literacia digital entre a população, ou mesmo a promoção de veículos independentes de apuração de fatos (AL SUR, 2021, p. 15).

4.6 CONCLUSÃO

O PL 2.630 constitui um importante esforço legislativo no sentido de proteger a liberdade de expressão na internet e promover um ambiente virtual saudável, responsável e que assegure o respeito às regras democráticas. É necessário, contudo, aprofundar o debate sobre o texto aprovado no Senado Federal, sanando problemas de técnica legislativa como enunciados confusos e provisões mal construídas que podem gerar problemas no futuro.

Entre um modelo de “regulação autoritária”, que obrigue as plataformas a monitorar e controlar de forma acintosa o conteúdo publicado por usuários, e um modelo de “autorregulação corporativa”, que não contrabalanceie o poder das empresas, deve-se buscar um caminho intermediário.

É preciso ponderar a busca pela proteção à liberdade de expressão com a preservação de bens caros à coletividade e ao interesse público. Nessa equação, o respeito aos direitos humanos deve desempenhar um papel central no desenho regulatório, garantindo o respeito às regras do jogo democrático e às instituições.

Para tal, a perspectiva procedimental de regulação da liberdade de expressão constitui um caminho possível para estabelecer instrumentos concretos para evitar violações a direitos fundamentais dos usuários na moderação privada de conteúdo.

Reguladores, usuários e plataformas de conteúdo digital devem trabalhar juntas para encontrar soluções legais, contratuais e técnicas que equilibrem a liberdade de expressão na internet com outros bens e valores constitucionais.

5 CONSIDERAÇÕES FINAIS

A governança de conteúdo online é um tema caro ao funcionamento das instituições democráticas e sensível ao exercício de direitos fundamentais dos brasileiros. Em 2021, foram contabilizados mais de 152 milhões de usuários de internet no Brasil, total que corresponde a quase 81% da população nacional com 10 anos ou mais (CETIC, 2021). Essas pessoas utilizam a internet para se comunicar, fazer negócios, buscar informações, ler notícias, consumir conteúdo, entre outros. Dada a sua centralidade e abrangência, a internet hoje constitui um verdadeiro pilar da vida social e democrática do país.

Todavia, esse ‘espaço público’ cibernético é, em grande medida, “governado” pelas plataformas digitais, que adotam termos de serviços e aplicam políticas de moderação próprias, intervindo de formas relevantes no conteúdo gerado por seus usuários. Essa dinâmica da governança privada do ambiente virtual tem chamado a atenção de governos e parlamentos ao redor do mundo, interessados em regular esses serviços, sobretudo as redes sociais. No entanto, boa parte das iniciativas legais experimentadas apresentam soluções desproporcionais, ou capazes de ampliar o risco de violações à liberdade de expressão. Enquanto alguns marcos legais estabelecem restrições quase proibitivas à moderação privada de conteúdo (ex.: MP 1068), outros delegam responsabilidades e obrigações públicas às plataformas, transformando-as em verdadeiros juízes ou polícias privadas dos conteúdos de terceiros na internet (ex.: NetzDG) (PALLERO *et al*, 2020, p. 7).

No Brasil, o Senado Federal aprovou em junho de 2020 o Projeto de Lei nº 2.630, que busca instituir a Lei Brasileira de Liberdade, Responsabilidade e Transparência na Internet, popularmente chamada de “PL das Fake News”. Essa proposta legislativa dispõe, entre outras coisas, sobre a moderação de conteúdo nas redes sociais. O texto seguiu para a Câmara dos Deputados, que tem agora a missão de analisar e deliberar sobre a proposta legislativa.

Face a esse contexto, o presente trabalho estabeleceu como *objetivo geral* descobrir se as inovações regulatórias de caráter procedimental previstas no PL 2.630/20 constituem uma resposta normativa robusta aos desafios da liberdade de expressão no século XXI, tomando como base a *perspectiva procedimental* de regulação das grandes plataformas digitais. Para atingir tal meta, foram constituídos

os seguintes *objetivos específicos*: (i) identificar os fundamentos e elementos constitutivos da abordagem regulatória procedimental, (ii) individualizar as disposições legais pertinentes previstas na proposta legislativa em questão, e, finalmente, (iii) analisar tais dispositivos à luz da perspectiva regulatória previamente delimitada.

Para tanto, buscou-se inicialmente delimitar o conceito e a extensão do direito fundamental à liberdade de expressão no ordenamento jurídico pátrio, explorando seus limites legais e aspectos relevantes à regulação das plataformas digitais, como o regime de responsabilização civil dos intermediários. Em seguida, foram descritas as transformações sofridas pelas abordagens regulatórias da liberdade de expressão (“*old school*” e “*new school of speech regulation*”) com a chegada da internet a partir das ideias de Jack Balkin. Logo após, foi apresentado um modelo pluralista para a regulação da liberdade de expressão proposto pelo mesmo autor, que descreve o cenário e os personagens dessa nova dinâmica: (i) as empresas de infraestrutura de internet, (ii) os Estados-nação e (iii) os emissores (“*speakers*”), sujeitos ativos do processo comunicativo (i.e., as três pontas do triângulo). Do jogo de forças entre esses grupos, surgem problemas como a censura prévia e colateral, uma governança privada de conteúdo pouco transparente, e sem garantias de devido processo, além da vigilância e manipulação dos usuários. Para enfrentar esses desafios, o professor norte-americano propõe, entre outras coisas, que sejam reconhecidas novas obrigações éticas, profissionais e legais às plataformas digitais, incluindo direitos típicos ao devido processo legal na moderação de conteúdo.

Com o cenário, atores e contexto bem delimitados, Bowers e Zittrain contribuem com o enredo da trama ao descrever a evolução na regulação da governança de conteúdo online, oferecendo um olhar único sobre o atual choque de valores e compreensões sobre a liberdade de expressão nas redes. Enquanto na Era do Direito, que se estendeu dos anos 90 aos anos 2010, buscou-se proteger a esfera de liberdade na internet, a Era da Saúde Pública, que se iniciou nos anos 2010 e segue até os dias de hoje, concentra-se em enfrentar os males do ambiente virtual, como a desinformação e os discursos de ódio, ainda que para isso seja necessário restringir a liberdade de expressão. Para superar o impasse entre esses dois conjuntos de valores, os autores propõem que seja construído um amplo consenso ao redor dos

processos decisórios sobre conteúdo, concentrando-se na busca por legitimidade processual, ao invés da definição de juízos substanciais sobre conteúdo. Esta proposta, ainda a ser implementada, caracteriza a Era do Processo.

A partir dos consensos e sugestões dos autores de referência, e com base em recomendações internacionais como os Princípios de Santa Clara (2018), e o documento “*Padrões para uma regulação democrática das grandes plataformas que garanta a liberdade de expressão online e uma Internet livre e aberta*” (2020), foram dispostos os elementos constitutivos da abordagem regulatória denominada “*perspectiva procedimental*”. Em suma, propugna-se o reconhecimento de novos direitos dos usuários face às plataformas digitais, e a adoção de mecanismos típicos do devido processo na moderação de conteúdo (ex.: direito ao recurso, direito à notificação e a decisões fundamentadas etc.). Também se defende a necessidade de uma autoridade com garantias de independência, autonomia técnica e decisória, com o distanciamento necessário para avaliar casos concretos, e garantir o cumprimento de suas decisões.

Tendo satisfeito o *primeiro objetivo específico*, partiu-se, então, no capítulo seguinte, para a descrição das provisões do PL 2.630/20. Dado o recorte metodológico focado no *processo de tomada de decisões* na governança de conteúdo online, a pesquisa documental concentrou-se em coletar apenas os dispositivos que dissessem respeito a tais *procedimentos*, seja na moderação de conteúdo em si, ou em relação aos *processos* globais de definição de regras e políticas. Em resumo, a proposta legislativa prevê mecanismos de transparência e devido processo além de garantias de ampla defesa e contraditório na curadoria de conteúdo, bem como a criação de um conselho consultivo multissetorial sobre o tema vinculado ao Senado Federal, e sugerir que as empresas constituam uma instituição de autorregulação regulada. Com os resultados da pesquisa documental devidamente discriminados, alcançou-se o *segundo objetivo específico*.

No último capítulo, por sua vez, foram analisadas as propostas do legislador brasileiro previamente delimitadas à luz da *perspectiva procedimental* de regulação das plataformas digitais. Concluiu-se, afinal, que conquanto seja bastante positiva a previsão de novos direitos e garantias processuais aos usuários brasileiros, como o direito ao recurso e à decisões fundamentadas, observou-se problemas de técnica

legislativa em alguns dispositivos (como o § 2º do art. 12), que poderão gerar interpretações ambíguas, e abrir espaço para arbitrariedades caso não sejam aperfeiçoados. Além disso, foram levantadas preocupações quanto à independência do Conselho de Transparência e Responsabilidade na Internet (CTRI), e os possíveis problemas de *enforcement* que um organismo de autorregulação regulada teria, não havendo garantias reais de que o interesse público seria priorizado em questões sensíveis à coletividade. É possível concluir, portanto, que todos os objetivos foram plenamente alcançados.

A hipótese central foi confirmada, indicando que o PL 2.630/20 acerta ao prever direitos e garantias dos usuários na moderação de conteúdo, mas que peca no modo como certos dispositivos foram redigidos (em especial o § 2º do art. 12), que podem dar azo a arbitrariedades na governança de conteúdo. Além disso, a previsão de um conselho multissetorial responsável por sugerir diretrizes para a moderação de conteúdo é salutar, mas a indicação de representantes de órgãos policiais é estranha ao propósito do colegiado, e pode abrir espaço para uma visão policial da moderação de conteúdo, capaz de colocar em risco direitos dos usuários. Ainda, a não previsão de um ente independente com poder normativo, não oferece garantias reais de que as recomendações serão seguidas pelas corporações da internet, o que pode dificultar a concretização de direitos e liberdades individuais.

A pergunta de pesquisa buscou responder se os mecanismos regulatórios de natureza procedimental previstos no PL 2.630/20 incorporam adequadamente os objetivos e recomendações da perspectiva procedimental. Após a análise de dados e a coleta de informações, concluiu-se que o marco regulatório proposto incorpora em parte as ideias da abordagem procedimental, mas falha em dar efetividade e um desenho regulatório claro e bem ajustado para a governança de conteúdo online.

A metodologia adotada caracterizou-se por uma pesquisa de finalidade básica estratégica, com objetivos descritivo e exploratório, realizada pelo método hipotético-dedutivo, com abordagem qualitativa e executada por meio de levantamento bibliográfico e documental. A estrutura do trabalho contemplou a concretização desse arcabouço metodológico, estando o desenvolvimento do trabalho delimitado em três capítulos bem definidos, sendo o primeiro dedicado ao referencial teórico, o segundo

voltado ao avanço da coleta de dados e o terceiro contemplando a apresentação e discussão dos resultados.

Não foi possível, contudo, analisar uma versão definitiva da proposta legislativa, que atualmente segue em tramitação pela Câmara dos Deputados. Desse modo, é bastante provável, e recomendável que os dispositivos aqui analisados sofram modificações pelo parlamento brasileiro, o qual ainda deve ajustar o desenho regulatório inicialmente aprovado. Portanto, análises posteriores serão bem-vindas e podem contribuir para a construção de marco regulatório mais robusto.

Certamente, a *perspectiva procedimental* adotada neste trabalho não pretende oferecer uma resposta a todos os desafios da governança de conteúdo online. Outras perspectivas regulatórias são certamente necessárias para lidar com as diversas facetas desse tema, como a questão concorrencial e o desafio da proteção de dados pessoais. Essa perspectiva regulatória, no entanto, oferece um caminho possível para proteger direitos e liberdades fundamentais, equilibrando as assimetrias de poder e informação, e protegendo a Internet como um ambiente livre, democrático e aberto.

6 REFERÊNCIAS

ALEMANHA. **Netzwerkdurchsetzungsgesetz**. Berlim. 01 set. 2017. Disponível em: <<https://www.gesetze-im-internet.de/netzdg/NetzDG.pdf>> Acesso em 10 de outubro de 2021.

ALVES, J. A. Lindgren. **A arquitetura internacional dos Direitos humanos**. São Paulo: Editora FTD, 1997, 186 p.

AL SUR. **Olhando Al Sur: Rumo a novos consensos regionais em matéria de responsabilidade de intermediários e moderação de conteúdos na Internet**. Abr. 2021. Disponível em: <<https://www.alsur.lat/pt-br/relatorio/olhando-al-sur-rumo-novos-consensos-regionais-em-materica-responsabilidade-intermediarios>>. Acesso em: 03 set. 2021.

AGAPITO, Leonardo; OLIVEIRA, José; MIRANDA, Matheus. **O modelo de “autorregulação regulada” e a teoria da captura: obstáculos à efetividade no combate à lavagem de dinheiro no Brasil**. 2017. REVISTA QUAESTIO IURIS. DOI: 10.12957/rqi.2017.26847

AMADO, Guilherme. **YouTube remove 14 lives de Bolsonaro; canal está perto de ser derrubado**. Metrôpoles. 21 jul. 2021. Brasília. Disponível em: <https://www.metropoles.com/colunas/guilherme-amado/youtube-remove-14-lives-de-bolsonaro-canal-esta-perto-de-ser-derrubado>. Acesso em: 23 out. 2021.

AMORIM, F.; ALVES, R.; MAIA, Flávia. **Pacheco devolve MP de Bolsonaro que alterava Marco Civil da Internet**. 14 set. 2021. Disponível em: <<https://www.jota.info/coberturas-especiais/liberdade-de-expressao/pacheco-devolve-mp-de-bolsonaro-que-alterava-marco-civil-da-internet-14092021>>. Acesso em: 25 out. 2021.

ARANHA, Márcio Iorio. **Manual de Direito Regulatório: Fundamentos de Direito Regulatório**. 3ª Edição. London: Laccademia Publishing, 2015.

_____. **As formas de autorregulação**. JOTA, Brasília, 26 de outubro de 2019. Disponível em: <<https://www.jota.info/opiniao-e-analise/artigos/as-formas-de-autorregulacao-26102019>>. Acesso em 15 de outubro de 2021.

ARCHEGAS, João Victor. ESTARQUE, Marina. **Redes Sociais e Moderação de Conteúdo: criando regras para o debate público a partir da esfera privada**. Abr. 2021. ITS Rio. Disponível em: <<https://itsrio.org/pt/publicacoes/redes-sociais-e-moderacao-de-conteudo/>>. Acesso em: 30 set. 2021.

ARRUDA, Vinicius Cervantes G. **Dados Pessoais e o Cumprimento da Obrigação Legal Prevista no Marco Civil da Internet**. Ricci Propriedade Intelectual. Jan. 2021. Disponível em: < <https://riccipi.com.br/dados-pessoais-e-o-cumprimento-da-obrigacao-legal-prevista-no-marco-civil-da-internet/>>. Acesso em 18 de outubro de 2021.

BALKIN, Jack M. **Some Realism About Pluralism: Legal Realist Approaches to the First Amendment**. Duke Law Journal, v. 1990, n. 3, pp. 375- 430, June 1990.

_____. **Free Speech and Hostile Environments**. Columbia Law Review, v. 99, pp. 1-28, 1999.

_____. **Digital Speech and Democratic Culture: a Theory of Freedom of Expression for the Information Society**. New York University Law Review, v.1, n.79, 2004.

_____. **The Future of Free Expression in a Digital Age**. Pepperdine Law Review, v. 36, 2009.

_____. **Free Speech is a Triangle**. Columbia Law Review, v. 118, p. 2011-2056, 2018.

_____. **Free speech in the algorithmic society: big data, private governance and, new school speech regulation**. UC Davis Law Review, v. 51, p. 1149-1210, 2018.

_____. **The First Amendment in the Second Gilded Age**. Buffalo Law Review, v. 66, n. 5, December 2018.

_____. **Fixing Social Media's Grand Bargain**. Hoover Working Group on National Security Technology, and Law, Aegis Paper Series, n. 1814, October 2018.

_____. **The Political Economy of Freedom of Speech in the Second Gilded Age**. Law and Political Economy blog, 4 de Julho de 2018. Disponível em: <<https://lpeblog.org/2018/07/04/the-political-economy-of-freedom-of-speech-in-the-second-gilded-age/>> Acesso em 15 de julho de 2019.

BARROSO, Luís Roberto. **Colisão Entre Liberdade de Expressão e Direitos da Personalidade**. Critérios de Ponderação. Interpretação Constitucionalmente Adequada do Código Civil e da Lei de Imprensa. Revista de Direito Administrativo, Rio de Janeiro, v. 235, pp. 1-36, Jan./Mar. 2004, p. 18.

_____. **Da caverna à internet: evolução e desafios da liberdade de expressão**. Revista Publicum. Rio de Janeiro, v. 6, n. 1, p. 1-12, 2020. Disponível em: <https://www.e-publicacoes.uerj.br/index.php/publicum/article/view/57576>>. Acesso em: 15 out. 2021.

BARLOW, John Perry. **A Declaration of the Independence of Cyberspace**, Electronic Frontier Foundation. 8 Fev. 1996. Disponível em: <<https://www.eff.org/cyberspace-independence>>. Acesso em 18 de outubro de 2021.

BBC. **6 controvérsias e escândalos do Facebook antes de mudar o nome para Meta**. 28 out. 2021. Disponível em: <https://www.bbc.com/portuguese/geral-59085894?at_custom3=BBC+Brasil&at_custom4=A7290838-38B4-11EC-9F64-4B240EDC252D&at_custom2=twitter&at_custom1=%5Bpost+type%5D&at_medium=custom7&at_campaign=64>. Acesso em 28 de outubro de 2021.

BERGMAYER, B. J. **Even Under Kind Masters: A Proposal to Require that Dominant Platforms Accord Their Users Due Process**. Mai. 2018. Public Knowledge

BRASIL. **Decreto nº 592, de 6 de julho de 1992**. Pacto Internacional sobre Direitos Civis e Políticos. Disponível em: <http://www.planalto.gov.br/ccivil_03/decreto/1990-1994/d0592.htm>. Acesso em 18 de outubro de 2021.

_____. **Constituição da República Federativa do Brasil de 1988**. Disponível em: <http://www.planalto.gov.br/ccivil_03/constituicao/constituicaocompilado.htm>. Acesso em: 22 set. 2021.

_____. **HC nº 82.424/RS**. Supremo Tribunal Federal. Relator: Min. Maurício Corrêa, Plenário, julgamento concluído em 19 de setembro de 2003.

_____. **Lei nº 12.965, de 23 de abril de 2014**. Disponível em: <http://www.planalto.gov.br/ccivil_03/ato2011-2014/2014/lei/l12965.htm>. Acesso em 18 de outubro de 2021.

_____. **ADI nº 4.451**. Supremo Tribunal Federal. Relator: Min. Alexandre de Moraes, PLENÁRIO, j. 21/06/2018, DJe 06/03/2019

_____. **Proposta de Emenda à Constituição nº 17, de 2019**. Senado Federal. Acrescenta o inciso XII-A, ao art. 5º, e o inciso XXX, ao art. 22, da Constituição Federal para incluir a proteção de dados pessoais entre os direitos fundamentais do cidadão e fixar a competência privativa da União para legislar sobre a matéria. Brasília: Senado Federal, 2021. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/135594>. Acesso em: 22 set. 2021.

_____. **Projeto de Lei nº 2.630. 2020**. Texto do PL 2630 aprovado no Senado e encaminhado à Câmara dos deputados. Disponível em: <https://www.camara.leg.br/proposicoesWeb/prop_mostrarintegra?codteor=1909983&filename=PL+2630/2020>. Acesso em: 22 set. 2021.

_____. **Parecer do Senador Angelo Coronel sobre o PL 26360**. Senado Federal. Disponível em: <<https://legis.senado.leg.br/sdleg-getter/documento?dm=8127649&ts=1630418526626&disposition=inline>>. Acesso em: 22 set. 2021.

_____. **Medida provisória nº 1.068, de 6 de setembro de 2021**. 2021a. Diário Oficial [da] República Federativa do Brasil, Poder Executivo, Brasília, DF. Publicado em: 06 set. 2021. Ed. 169-A. Seç: 1 - Extra A. Pág. 1.

_____. **Medida Provisória reforça direitos e garantias de usuários de redes sociais**. Secretaria Especial de Comunicação Social da Presidência da República (SECOM). 2021b. Disponível em: <https://www.gov.br/planalto/pt-br/acompanhe-o-planalto/noticias/2021/09/medida-provisoria-reforca-direitos-e-garantias-de-usuarios-de-redes-sociais>

_____. **Ato Declaratório Do Presidente Da Mesa Do Congresso Nacional Pela Devolução Da MP 1068.** 2021c. Disponível em: <<https://drive.google.com/file/d/1mEObfZCDPUweV9ysZyl5nZxeJFqecqk-/view>>. Acesso em: 22 set. 2021.

BOWERS, John; ZITTRAIN, Jonathan. **Answering Impossible Questions: Content Governance in an Age of Disinformation.** 2020. Harvard Kennedy School Misinformation Review, 1–11. Disponível em: <<https://doi.org/10.37016/mr-2020-005>>. Acesso em: 02 abr. 2021.

CABETTE, Eduardo. **Classificação Doutrinária da Censura.** Jusbrasil. 2015. Disponível em: <<https://eduardocabette.jusbrasil.com.br/artigos/236728638/classificacao-doutrinaria-da-censura>>. Acesso em: 22 set. 2021.

CANAZZA, Mario Rodrigo. **The Internet as a global public good and the role of governments and multilateral organizations in global internet governance.** 2018. Meridiano 47. JOURNAL OF GLOBAL STUDIES. Disponível em: <<http://dx.doi.org/10.20889/M47e19007>>. Acesso em: 17 out. 2021.

CANOTILHO, J. J. Gomes. **Direito constitucional.** 6. ed. Coimbra: Livraria Almedina, 1995.

CDR - COALIZÃO DIREITOS NA REDE. **PL das Fake News: Identificação e Criminalização em massa de usuários.** 20 jun. 2020. Disponível em: <<https://direitosnarede.org.br/2020/06/20/pl-das-fake-news-identificacao-e-criminalizacao-em-massa-de-usuarios/>>. Acesso em: 22 set. 2021.

CELESTE, Edoardo. **Digital constitutionalism: a new systematic theorization.** 03 jan. 2019. Disponível em: <<https://doi.org/10.1080/13600869.2019.1562604>>. Acesso em: 22 set. 2021.

CEROY, Frederico Meinberg. **Os conceitos de provedores no Marco Civil da Internet.** Revista Jus Navigandi, ISSN 1518-4862, Teresina, ano 19, n. 4093, 15 set. 2014. Disponível em: <<https://www.migalhas.com.br/depeso/211753/os-conceitos-de-provedores-no-marco-civil-da-internet>>. Acesso em: 22 set. 2021.

CETIC - Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação. **Cresce o uso de Internet durante a pandemia e número de usuários no Brasil chega a 152 milhões, é o que aponta pesquisa do Cetic.br.** 18 ago. 2021. Disponível em: <[https://cetic.br/pt/noticia/cresce-o-uso-de-internet-durante-a-pandemia-e-numero-de-usuarios-no-brasil-chega-a-152-milhoes-e-o-que-aponta-pesquisa-do-cetic-br/#:~:text=O%20Brasil%20tem%20152%20milh%C3%B5es,com%2010%20anos%20ou%20mais.&text=Pela%20primeira%20vez%2C%20o%20levantamento,que%20in div%C3%ADduos%20usu%C3%A1rios%20\(81%25\)](https://cetic.br/pt/noticia/cresce-o-uso-de-internet-durante-a-pandemia-e-numero-de-usuarios-no-brasil-chega-a-152-milhoes-e-o-que-aponta-pesquisa-do-cetic-br/#:~:text=O%20Brasil%20tem%20152%20milh%C3%B5es,com%2010%20anos%20ou%20mais.&text=Pela%20primeira%20vez%2C%20o%20levantamento,que%20in div%C3%ADduos%20usu%C3%A1rios%20(81%25))>. Acesso em: 23 set. 2021.

CGI - COMITÊ GESTOR DA INTERNET NO BRASIL. **Relatório Internet, Desinformação e Democracia.** 17 mar. 2020. Disponível em:

<<https://www.cgi.br/publicacao/relatorio-internet-desinformacao-e-democracia/>>. Acesso em: 19 out. 2021.

CHEDE, Cezar T. **Padrões abertos, interoperabilidade e interesse público**. PoliTICS. Nov 2008. Disponível em: <<https://politics.org.br/edicoes/padr%C3%B5es-abertos-interoperabilidade-e-interesse-p%C3%ABlico>>. Acesso em: 02 out. 2021.

CIDH – Comissão Interamericana de Direitos Humanos. **Liberdade de Expressão na Internet. Relatoria Especial para a Liberdade de Expressão**. 2013. Disponível em: <http://www.oas.org/pt/cidh/expressao/docs/publicaciones/2014%2008%2004%20liberdade%20de%20express%C3%A3o%20e%20internet%20rev%20%20hr_rev%20lar.pdf>. Acesso em 13 out. 2021.

CITRON, Danielle; WITTES, Benjamin. **The Internet Will Not Break: Denying Bad Samaritans § 230 Immunity**. 2017. 86 Fordham L. Rev. 401. Disponível em: <<https://ir.lawnet.fordham.edu/flr/vol86/iss2/3>>. Acesso em: 15 set. 2021

DICTIONARY.COM. **Disinformation**. 2021. Disponível em: <<https://www.dictionary.com/browse/disinformation>>. Acesso em 12 out. 2021.

DOUEK, Evelyn. **Verified accountability: Self-regulation of content moderation as an answer to the special problems of speech regulation**. 18 set. 2019. Lawfare. Disponível em: <<https://www.lawfareblog.com/verified-accountability-self-regulation-content-moderation-answer-special-problems-speech-0>>. Acesso em: 24 abr. 2021.

EC. European Commission. **A multi-dimensional approach to disinformation**. 2018. High Level Group on Fake News and Online Disinformation. Disponível em: <<https://op.europa.eu/et/publication-detail/-/publication/6ef4df8b-4cea-11e8-be1d-01aa75ed71a1/>>. Acesso em: 19 out. 2021.

ELLUL, Jacques. **The Technological Society**. 1954. New York: Knopf, 1967. Print. P. 429

EFF - ELECTRONIC FRONTIER FOUNDATION; *et al.* **Princípios de Manila Sobre Responsabilidade dos Intermediários**. 2015. Disponível em: <<https://manilaprinciples.org/pt-br.html#:~:text=N%C3%A3o%20se%20deve%20solicitar%20a,autoridade%20judicial%20independente%20e%20imparcial>>. Acesso em: 20 set. 2021.

_____. **THE SANTA CLARA PRINCIPLES On Transparency and Accountability in Content Moderation**. 2018. Disponível em: <<https://santaclaraprinciples.org/>>. Acesso em: 15 jul. 2021.

EUA - ESTADOS UNIDOS DA AMÉRICA. **U.S. Code § 230 - Protection for private blocking and screening of offensive material**. 1996. Disponível em: <<https://www.law.cornell.edu/uscode/text/47/230>>. Acesso em: 19 out. 2021.

_____. **Packingham v. North Carolina**, 137 S. Ct. 1730. Supreme Court of the United States. 2017. Sandvig v. Sessions, 315 F. Supp. 3d 1, 12 (D.D.C. 2018).

Disponível em: <https://www.supremecourt.gov/opinions/16pdf/15-1194_0811.pdf>. Acesso em 28 out. 2021.

FACEBOOK. **Stats. Facebook Newsroom.** 2018. Disponível em: <<https://about.facebook.com/company-info/#statistics>>. Acesso em: 23 set. 2021.

FARIAS, Edilsom P. **Liberdade de expressão e comunicação: teoria e proteção constitucional.** Tese de doutorado. 2001. Universidade Federal de Santa Catarina, Centro de Ciências Jurídicas. Programa de Pós-Graduação em Direito, 2001. Disponível em: <<https://repositorio.ufsc.br/handle/123456789/79426>>. Acesso em: 19 out. 2021.

FONSECA, Gabriel Campos Soares da. **O Mercado de Ideias: Liberdade de Expressão, Plataformas Digitais e Regulação da Internet.** Monografia de Final de Curso (Graduação em Direito), Faculdade de Direito, Universidade de Brasília, Brasília, DF, 105f, 2019. Disponível em: <<https://bdm.unb.br/handle/10483/25489>>. Acesso em: 25 mai. 2021.

FRANÇA. **Creating a French framework to make social media platforms more accountable: Acting in France with a European vision.** Mai. 2019. Disponível em: <<https://thecre.com/RegSM/wp-content/uploads/2019/05/French-Framework-for-Social-Media-Platforms.pdf>>. Acesso em 18 out. 2021.

FRAZÃO, Ana; MEDEIROS, Ana Rafaela. **Responsabilidade civil dos provedores de internet: a liberdade de expressão e o art. 19 do Marco Civil.** 23 fev. 2021. Disponível em: <<https://www.migalhas.com.br/coluna/migalhas-de-responsabilidade-civil/340656/responsabilidade-civil-dos-provedores-de-internet>>. Acesso em: 19 out. 2021.

FRAZÃO, Ana. **Devido processo digital. Em que medida o devido processo legal se aplica aos julgamentos online e às relações privadas na internet?** JOTA. 20 out. 2021. Disponível em: <<https://www.jota.info/opiniao-e-analise/colunas/constituicao-empresa-e-mercado/devido-processo-digital-20102021?amp#ftnref8>>. Acesso em: 23 out. 2021.

_____. **Plataformas digitais e os desafios para a regulação jurídica.** In: PARENTONI, Leonardo. (Coord.) (et. al) *Direito, Tecnologia e Inovação.* Belo Horizonte: Editora D'Plácido, 2018.

GILLESPIE, Tarleton. **Custodians of the Internet: platforms, content moderation, and the hidden decisions that shape social media.** 2018. [Versão Eletrônica]. New Haven: Yale University Press.

IDEC - INSTITUTO BRASILEIRO DE DEFESA DO CONSUMIDOR. **Você sabe o que é “moderação de conteúdo” nas mídias digitais?** Seus direitos nas mídias digitais. 2021. Disponível em: <<https://idec.org.br/direitos-nas-midias>>. Acesso em: 18 out. 2021.

IRIS – Instituto de Referência em Internet e Sociedade. **Liberdade de expressão, moderação de conteúdo e o PL das fake news**. 06 jul. 2020. Disponível em: <<https://irisbh.com.br/liberdade-de-expressao-moderacao-de-conteudo-e-o-pl-das-fake-news/>>. Acesso em: 27 set. 2021.

KARPEN, Ulrich. **Freedom of expression**, in U. Karpen (ed.), *The Constitution of the Federal Republic of Germany*, Baden-Baden: Nomos Verlagsgesellschaft, 1988, p. 93

KAUL, Inge; CONCEIÇÃO, Pedro; LE GOULVEN, Katell; MENDOZA, Ronald U. **Providing Global Public Goods: Managing Globalization**. 2003. New York: Oxford University Press.

KLONICK, Kate. **The New Governors: The People, Rules, and Processes Governing Online Speech**. 6 abr. 2018. Harvard Law Review, vol. 131. Disponível em: <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2937985>. Acesso em: 17 jun. 2021.

KOSEFF, Jeff. **The Twenty-Six Words That Created The Internet**. Cornell Univ. Press, 2019

LAFER, Celso. **Parecer – o caso Ellwanger: anti-semitismo como crime da prática do racismo**. Revista de informação legislativa, [S. l.], v. 41, n. 162, p. 53-89, abr./jun. de 2004. Disponível em: <<http://www2.senado.leg.br/bdsf/bitstream/handle/id/948/R162-08.pdf?sequence=4>>. Acesso em: 20 out. 2021.

LANE, Tyler. **The Public Forum Doctrine in the Modern Public Square**. 2019. Ohio Northern University Law Review, vol. 45, no. 2, 2019, p. 465-viii.

LAPIN - LABORATÓRIO DE POLÍTICAS PÚBLICAS E INTERNET. **Isso é fake news?** Dez. 2020. Disponível em: https://lapin.org.br/wp-content/uploads/2020/10/Isso-e-%CC%81-Fake-News-Um-guia-ra-%CC%81pido-de-desinformac-%CC%A7a-%CC%83o-na-internet_LAPIN.pdf. Acesso em: 20 set. 2021.

LASLO, Matt. **The Fight Over Section 230 – and the Internet as We Know It**. 13 Aug. 2019. Wired. Disponível em: <<https://www.wired.com/story/fight-over-section-230-internet-as-we-know-it/>>. Acesso em: 21 out. 2021.

MARTINELLI, Gustavo. **Os limites e deveres da liberdade de imprensa**. Aurum. 17 jul. 2020. Disponível em: <<https://www.aurum.com.br/blog/liberdade-de-imprensa/>>. Acesso em: 20 set. 2021.

MENA, Isabela. **Verbete Draft: o que é Techlash**. 26 fev. 2020. Draft. Disponível em: <<https://www.projtodraft.com/verbete-draft-o-que-e-techlash/>>. Acesso em: 20 set. 2021.

MENDES, Gilmar, Ferreira; BRANCO, Paulo Gustavo Gonet. **Curso de Direito Constitucional**. 10. ed. rev. e atual. São Paulo: Saraiva, 2015.

MONTEIRO, Artur Pericles Lima. **Armadilhas à liberdade de expressão na MP 1068/2021**. 14 set. 2021. Disponível em: <<https://www.jota.info/stf/supra/armadilhas-a-liberdade-de-expressao-na-mp-1068-2021-14092021>>. Acesso em: 20 set. 2021.

MONTEIRO, Artur Pericles Lima; CRUZ, Francisco Brito; SILVEIRA, Juliana Fonteles; VALENTE, Mariana G. **Armadilhas e caminhos na regulação da moderação de conteúdo. Diagnósticos & Recomendações**. São Paulo: InternetLab, 2021. Disponível em: <https://www.internetlab.org.br/wp-content/uploads/2021/09/internetlab_armadilhas-caminho-moderacao.pdf>. Acesso em: 20 set. 2021.

MOSTERT, F. **“Digital Due Process”: A Need for Online Justice**. Jan. 2020. Journal of Intellectual Property Law & Practice. DOI: 10.1093/jiplp/jpaa024.

NEVES, Daniel Amorim Assumpção. **Manual de direito processual civil**. 2018. Vol. único. 10. Ed. JusPODIVM. ISBN: 9788544219447.

OLIVA, Thiago. **Responsabilidade de intermediários e a garantia da liberdade de expressão na rede**. 23 abr. 2019. InternetLab. Disponível em: <<https://www.internetlab.org.br/pt/especial/responsabilidade-de-intermediarios-e-a-garantia-da-liberdade-de-expressao-na-rede/>>. Acesso em: 28 set. 2021.

OLIVIERI, A. C. **Censura - O regime militar e a liberdade de expressão**. História do Brasil. UOL. Mar. 2014. Disponível em: <<https://educacao.uol.com.br/disciplinas/historia-brasil/censura-o-regime-militar-e-a-liberdade-de-expressao.htm?cmpid=copiaecola>>. Acesso em: 18 out. 2021.

ONU – ORGANIZAÇÃO DAS NAÇÕES UNIDAS. **Declaração Universal dos Direitos Humanos**. 1948. Disponível em: <https://www.ohchr.org/en/udhr/documents/udhr_translations/por.pdf>. Acesso em: 18 set. 2021.

_____. **Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression**. A/HRC/38/35. Human Rights Council. 06 de abril de 2018. Disponível em: <<https://www.ohchr.org/EN/Issues/FreedomOpinion/Pages/ContentRegulation.aspx>>. Acesso em: 22 set. 2021.

_____. **Report of the Special Representative of the Secretary General on the issue of human rights and transnational corporations and other business enterprises, John Ruggie**. CCPR/C/GC/34. Human Rights Council. Mar./ 2011. Disponível em: <https://www.ohchr.org/documents/issues/business/a-hrc-17-31_aev.pdf>. Acesso em: 22 set. 2021.

OBSERVACOM. **Fuerte rechazo a proyecto de ley sobre Fake News en Brasil: “la peor ley de Internet del mundo”**. 23 jun. 2020. Disponível em: <<https://www.observacom.org/fuerte-rechazo-a-proyecto-de-ley-sobre-fake-news-en-brasil-la-peor-ley-de-internet-del-mundo/>>. Acesso em: 22 set. 2021.

ORRICO, Alexandre; SPAGNUOLO, Sérgio; MARTINS, Laís. **MEDIDA BOLSONARO SOBRE MODERAÇÃO DE REDES PEGOU MAL COM TODO MUNDO**. 09 set. 2021. Núcleo Jornalismo. Disponível em: <<https://www.nucleo.jor.br/reportagem/2021-09-09-redes-medida-bolsonaro-moderacao/>>. Acesso em: 22 set. 2021.

PALLERO, Javier; BARATA, Joan; PIAZZA, Andrés; MASTRINI, Guillermo; BECERRA, Martín. FREULER, Juan Ortiz. **Padrões para uma regulação democrática das grandes plataformas que garanta a liberdade de expressão online e uma Internet livre e aberta**. Jul. 2020. OBSERVACON, et al. Disponível em: <<https://www.observacom.org/wp-content/uploads/2020/09/Padroes-para-uma-regulacao-democratica-das-grandes-plataformas.pdf>>. Acesso em: 22 set. 2021.

POPPER, Karl Raimund. **A sociedade aberta e seus inimigos**. 1902. Tradução de Milton Amado. Belo Horizonte, Ed. Itatiaia; São Paulo, Ed. da Universidade de São Paulo, 1974.

REALE JÚNIOR. **Limites à liberdade de expressão**. Revista Espaço Jurídico, Florianópolis, v. 11, n. 2, p.; 374-401, jul./dez. 2010.

ROBERTS, Sarah T. **Content Moderation**. In: Schintler L., McNeely C. (eds) Encyclopedia of Big Data. Springer, Cham. 2017. Disponível em: <https://doi.org/10.1007/978-3-319-32001-4_44-1>. Acesso em: 18 out. 2021.

RUEDIGER, M. A.; GRASSI, A. (Coord.). **Discurso de ódio em ambientes digitais: definições, especificidades e contexto da discriminação on-line no Brasil a partir do Twitter e do Facebook**. Policy paper. Rio de Janeiro: FGV DAPP, 2021.

SANTOS, Bruna Martins. **Uma avaliação do Modelo de Responsabilidade de Intermediários do Marco Civil para o desenvolvimento da Internet no Brasil**. Dez. 2020. Disponível em: <<https://politics.org.br/edicoes/uma-avalia%C3%A7%C3%A3o-do-modelo-de-responsabilidade-de-intermedi%C3%A1rios-do-marco-civil-para-o>>. Acesso em: 22 set. 2021.

SARLET, Ingo Wolfgang. **Direitos fundamentais, fake news e democracia: notas acerca da MP 1.068**. GenJurídico. 30 set. 2021. Disponível em: <<http://genjuridico.com.br/2021/09/30/direitos-fundamentais-mp-1068/>>. Acesso em: 18 out. 2021.

SAFERNET BRASIL. “[...] **A MP limita a autonomia das plataformas em desenvolver e aplicar suas políticas e tecnologias de detecção e combate a abusos online [...]**” 6 set. 2021. Disponível em: <<https://twitter.com/safernet/status/1435008124818280454>>. Acesso em: 18 out. 2021.

SENRA, Ricardo. **Após Twitter, Facebook e Instagram excluem vídeo de Bolsonaro por 'causar danos reais às pessoas'**. BBC. 30 mar. 2020. Disponível em: <<https://www.bbc.com/portuguese/brasil-52101240>>. Acesso em: 23 out. 2021.

SCHERTEL, Laura. **Decisão histórica do STF reconhece o direito fundamental à proteção de dados pessoais**. 10 mai. 2020. JOTA. Disponível em:

<<https://www.jota.info/opiniao-e-analise/artigos/decisao-historica-do-stf-reconhece-direito-fundamental-a-protecao-de-dados-pessoais-10052020>>. Acesso em: 23 set. 2021.

SOUZA, Carlos Affonso Pereira. **Análise: Inconstitucional, MP de Bolsonaro cria 'Ministério da Mentira'**. UOL, Tilt. Rio de Janeiro, 07 set. 2021. Disponível em: <<https://www.uol.com.br/tilt/colunas/carlos-affonso-de-souza/2021/09/07/medida-provisoria-de-bolsonaro-sobre-redes-sociais-e-inconstitucional.htm?cmpid=copiaecola>>. Acesso em: 25 out. 2021.

SOUZA, Carlos Affonso Pereira; TEFFÉ, Chiara Spadaccini. **Liberdade de Expressão e o Marco Civil da Internet**. Pesquisa TIC Domicílios - Artigos, 2016. Disponível em: <https://www.academia.edu/36006753/LIBERDADE_DE_EXPRESS%C3%83O_E_O_MAR%20CO_CIVIL_DA_INTERNET> Acesso em: 27 de set. de 2021.,

_____. **Responsabilidade dos provedores por conteúdos de terceiros na internet**. 23 jan. 2017. Consultor Jurídico. Disponível em: <<https://www.conjur.com.br/2017-jan-23/responsabilidade-provedor-conteudo-terceiro-internet>>. Acesso em 13 out. 2021.

_____. **Regulação de Plataformas Digitais: debates e perspectivas. 5ª aula. Regimes de Responsabilização das Plataformas**. Mai. 2021. ITS Rio. Disponível em: <<https://itsrio.org/pt/cursos/regulacao-de-plataformas-digitais-realidades-e-alternativas/>>. Acesso em: 03 mai. 2021.

STARZAK, Alissa. **When the Internet (Officially) Became the Public Square**. 21 jun. 2017. Cloudflare. Disponível em: <<https://blog.cloudflare.com/internet-became-public-square/>>. Acesso em: 27 out. 2021.

STATISA. **Regional distribution of desktop traffic to Reddit.com as of June 2021, by country**. Jun. 2021. Disponível em: <<https://www.statista.com/statistics/325144/reddit-global-active-user-distribution/>>. Acesso em: 12 out. 2021.

STERN, Marc A.; KAUL, Inge; GRUNBERG, Isabelle. **Global Public Goods: International Cooperation in the 21st Century**. 1999. New York: Oxford University Press.

STUCKE, Maurice E. **Should We Be Concerned About Data-Opolies?** 2018. Disponível em: <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3144045>>. Acesso em: 02 mai. 2020.

TÔRRES, Fernanda Carolina. **O direito fundamental à liberdade de expressão e sua extensão**. 2013. Revista de informação legislativa, v. 50, n. 200, p. 61-80, out./dez. 2013. Disponível em: <http://www2.senado.leg.br/bdsf/handle/id/502937>. Acesso em: 02 out. 2021.

UCHINAKA, Fabiana. **Bolsonaro muda Marco Civil da Internet para limitar remoção de posts.** UOL. 06 set. 2021. Disponível em: <<https://www.uol.com.br/tilt/noticias/redacao/2021/09/06/bolsonaro-muda-marco-civil-da-internet-e-proibe-que-rede-social-apague-post.htm>>. Acesso em: 23 set. 2021.

UK – UNITED KINGDOM. **Online Harms White Paper.** 8 Abr. 2019. Disponível em: <<https://www.gov.uk/government/consultations/online-harms-white-paper>>. Acesso em: 20 out. 2021.

_____. **Draft Online Safety Bill.** Mai. 2021. Disponível em: <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/985033/Draft_Online_Safety_Bill_Bookmarked.pdf>. Acesso em: 03 nov. 2021.

VALERIO, Nathalya. **Qual a diferença entre processo e procedimento?** 01 nov. 2019. MasterJuris. Disponível em: <<https://masterjuris.com.br/qual-a-diferenca-entre-processo-e-procedimento/>>. Acesso em 31 ago. 2021.

VERONESE, Alexandre; FONSECA, Gabriel. **Desinformação, fake news e mercado único digital: a potencial convergência das políticas públicas da União Europeia com os Estados Unidos para melhoria dos conteúdos comunicacionais.** Cadernos Adenauer, São Paulo, ano XIX, v. 4, 2018. Disponível em: <https://www.researchgate.net/publication/334811848_Desinformacao_fake_news_e_mercado_unico_digitala_potencial_convergencia_das_politicas_publicas_da_Uniao_Europeia_com_os_Estados_Unidos_para_melhoria_dos_conteudos_comunicacionais>. Acesso em: 27 set. 2020.

_____. **O Futuro da Democracia na Era Digital: uma análise exploratória do Marco Civil da Internet.** Iniciação Científica (Graduação em Direito), Universidade de Brasília, Brasília, 2017.

VIEIRA, Rodrigo. **Moderação de conteúdo na internet brasileira: em defesa do Marco Civil.** Consultor Jurídico. 09 set. 2021. Disponível em: <https://www.conjur.com.br/2021-set-09/rodrigo-vieira-moderacao-conteudo-internet-brasileira>. Acesso em: 23 set. 2021.

PEREIRA JR., Ademir Antonio; VIEIRA, Yan Villela. **MP 1.068, regulação de conteúdo em redes sociais e livre iniciativa.** 21 set. 2021. Consultor Jurídico. Disponível em: <<https://www.conjur.com.br/2021-set-21/opiniao-mp-1068-regulacao-conteudo-redes-sociais>>. Acesso em: 23 set. 2021.

VIOLANTE, João Luís Mousinho dos Santos Monteiro. **O caso *Ellwanger* e seu impacto no direito brasileiro.** 2010. Dissertação (Mestrado em Direito). Pontifícia Universidade Católica de São Paulo, São Paulo, 2010. Disponível em: <<https://tede2.pucsp.br/bitstream/handle/5345/1/Joao%20Luis%20Mousinho%20dos%20Santos%20Monteiro%20Violante.pdf>>. Acesso em: 20 out. 2021.

ZIMDARS, Melissa; MCLEOD, Kembrew. **Fake news: understanding media and misinformation in the digital age.** 2020. Cambridge, Massachusetts: The MIT Press.

ZITTRAIN, Jonathan. **Three Eras of Digital Governance**. 23 Set. 2019. Berkman Center for Internet & Society. Disponível em: <<https://dx.doi.org/10.2139/ssrn.3458435>>. Acesso em: 06 ago. 2021.

ZUCKERBERG, Mark. **A Blueprint for Content Governance and Enforcement**. Facebook. 15 nov. 2018. Disponível em: <<https://www.facebook.com/notes/751449002072082/>>. Acesso em: 02 out. 2021.

WU, Tim. **The Attention Merchants: The Epic Scramble to Get Inside Our Heads**. 21. New York, NY. 2016.

_____. **Subtle and Insidious, Technology Is Designed to Addict Us**, Wash. Washington Post. 2 Mar 2017. Disponível em: <<https://www.washingtonpost.com/opinions/subtle-and-insidious-technology-is-designed-to-addict-us/2017/03/02/5b983ef4-fcee-11e6-99b4-9e613afeb09f>>. Acesso em: 10 out. 2021.

_____. **Impérios da Comunicação: do telefone à internet, da AT&T ao Google**. Trad. Cláudio Carina. Rio de Janeiro: Zahar, 2012, p. 151-152.