

INSTITUTO BRASILEIRO DE ENSINO, DESENVOLVIMENTO E PESQUISA
ESCOLA DE DIREITO E ADMINISTRAÇÃO PÚBLICA
MESTRADO ACADÊMICO EM DIREITO CONSTITUCIONAL

SULDBLANO OLIVEIRA GOMES

LIBERDADE DE EXPRESSÃO E GOVERNANÇA DA *INTERNET*: remédios
alternativos da moderação de conteúdo nas plataformas digitais.

BRASÍLIA

2022

SULDBLANO OLIVEIRA GOMES

LIBERDADE DE EXPRESSÃO E GOVERNANÇA DA *INTERNET*: remédios
alternativos da moderação de conteúdo nas plataformas digitais.

Dissertação de Mestrado desenvolvida sob a
orientação do Prof. Dr. Ilton Norberto Robl Filho
apresentado para obtenção do Título de Mestre em
Direito Constitucional.

BRASÍLIA

2022

SULDBLANO OLIVEIRA GOMES

LIBERDADE DE EXPRESSÃO E GOVERNANÇA DA *INTERNET*: remédios
alternativos da moderação de conteúdo nas plataformas digitais.

Dissertação de Mestrado apresentada ao Programa
de Pós-Graduação *Stricto Sensu* em Direito
Constitucional do IDP, como requisito para obtenção
do título de Mestre em Direito Constitucional

Data: 20/12/2022, às 14h00.

BANCA EXAMINADORA

Prof. Dr. Ilton Norberto Robl Filho
Instituto Brasiliense de Direito Público – IDP

Prof. Dr. Paulo Gustavo Gonet Branco
Instituto Brasiliense de Direito Público – IDP

Prof. Dr. Gustavo Ferreira Santos
Universidade Católica de Pernambuco - UNICAP

AGRADECIMENTOS

Primeiramente, agradeço a Deus pelo dom da vida e pela força para encarar os desafios diários.

À minha amada mãe Fátima Oliveira, pela acolhida materna que sempre acalma a alma e por ser a guerreira dos atos heroicos de todos os dias. Além dos principais ensinamentos e exemplo de vida, por sempre insistir que devemos persistir em nossos sonhos.

Agradeço aos meus amados irmãos Sullivan Gomes, Sulleandro Gomes, Sulimary Gomes e Vinicio Oliveira, que mesmo na correria diária e com a distância, sempre emanaram energias positivas. Aos meus sobrinhos Victória Cristina Gomes, Lucas Ribeiro, Joffre Júnior Gomes e, o mais novo “mascote”, Théo Gomes pela doçura e entusiasmo de ver a vida com mais leveza.

Louvo e sou grato por toda atenção e paciência, ao meu querido mestre Melquíades Barroso, o pai que a vida me ofertou e, no alto de sua sapiência, me mostrou o mundo do conhecimento, da música e das artes.

Gratidão aos meus amigos de sempre, Daniel Soares Veloso, Rômulo Lima dos Santos, José Pereira Lima Filho e Sávio Ramon que mesmo com a distância nunca desistiram e desacreditaram de mim.

Ao amigo Odélio Garcia, pela compreensão e enorme auxílio nesses dias corridos, dividindo comigo as árduas atribuições Ministeriais.

Aos servidores do Ministério Público do Estado do Pará que me suportaram durante essa jornada dupla: Paula Geovana, Aline Rodrigues, Ana Paula Barroso, Carla Andressa, Valdir Teixeira, Ktlei Rayane e Luciana Oliveira.

Aos educadores do IDP, que se dispuseram a compartilhar o conhecimento. Em especial, agradeço ao Professor Dr. Ilton Norberto Robl Filho, por acreditar no potencial da pesquisa e por todo o incentivo necessário. Ao Professor Dr. Paulo Gonet, pela generosidade e apoio, além das importantíssimas contribuições durante o curso e a qualificação. Agradeço, também, o Professor Dr. Rodrigo Luís Kanayama pelas valiosas contribuições na qualificação. Por fim, ao Prof. Dr. Gustavo Ferreira Santos, pela disponibilidade e solidariedade acadêmica.

“Outro tempo. Outro direito.”

Cesare Vivante

RESUMO

O estudo possui como tema a liberdade de expressão no ecossistema digital sob a perspectiva da governança privada da moderação de conteúdo pelas plataformas digitais, com base em termos de uso e serviços próprios. O trabalho possui o seguinte problema de pesquisa: *É possível as redes sociais removerem conteúdos de usuários? Se sim, é o único meio hábil para corrigir violações as suas regras, e como forma de garantir o uso racional da liberdade de expressão nas redes sociais?* E o objetivo geral é *compreender as principais controvérsias sobre a operação da governança privada na moderação de conteúdo e como o direito brasileiro trata tal problemática*. Objetiva, especificamente, *fornecer* um panorama da liberdade de expressão no ambiente virtual, identificando novas ameaças e uma nova postura do Estado como garantidor deste direito fundamental; *Compreender* como opera a governança privada das plataformas digitais, analisando atores privados, sobretudo quanto às regras elaboradas pelas plataformas *online* e os direitos fundamentais, especialmente, a liberdade de expressão; Bem como, pretende *identificar* os principais tópicos de discussão sobre a moderação de conteúdo pelas plataformas *online*, apresentar rol de remédios para correção de conteúdos ou ações que violem as regras internas das empresas de mídia social para além da simples remoção de conteúdo. Trata-se de pesquisa bibliográfica e documental, de natureza qualitativa. A revisão da literatura realizada inclui a revisão da legislação; termos de uso e serviço de mídia social; declarações de direitos digitais e outros instrumentos internacionais; artigos científicos disponíveis em periódicos nacionais e internacionais; livros; monografias, dissertações e teses publicadas em repositórios institucionais de renomadas universidades brasileiras e estrangeiras. Além disso, foram analisados princípios e conceitos de direito internacional público e privado, bem como a aplicação horizontal dos direitos humanos e dos termos de serviço das plataformas de mídia social. O resultado da pesquisa sugere que a defesa da liberdade de expressão tomou proporções mais complexas, originando a necessidade de encontrar um arranjo de moderação de conteúdo que melhor amolde os aspectos práticos da fala *online*, devendo as plataformas de mídias sociais adotarem em suas normas internas um teor de direitos fundamentais, sobretudo quanto à liberdade de expressão, bem como avaliar cenários com soluções alternativas para moderação de conteúdo, encarando novos padrões de proporcionalidade e probabilidade a serem explorados, cujo objetivo seja a melhoria do processo de moderação de conteúdo.

PALAVRA-CHAVE: Direitos Fundamentais. Moderação de Conteúdo. Mídia Social. Remédios alternativos.

ABSTRACT

The study has as its theme freedom of expression in the digital ecosystem from the perspective of private governance of content moderation by digital platforms, based on terms of use and services. The work has the following research problem: Is it possible for social networks to remove user content? If so, is it the only skillful way to correct violations of its rules, and as a way to guarantee the rational use of freedom of expression in social networks? And the general objective is to understand the main controversies about the operation of private governance in content moderation and how Brazilian law deals with this problem. It aims, specifically, to provide an overview of freedom of expression in the virtual environment, identifying new threats and a new posture of the State as guarantor of this fundamental right; Understand how the private governance of digital platforms operates, analyzing private actors, especially regarding the rules developed by online platforms and fundamental rights, especially freedom of expression; As well as, it intends to identify the main topics of discussion about the moderation of content by online platforms, to present a list of remedies for correcting content or actions that violate the internal rules of social media companies beyond the simple removal of content. This is a bibliographical and documentary research, of a qualitative nature. The literature review carried out includes the review of legislation; terms of use and social media service; digital rights bills and other international instruments; scientific articles available in national and international journals; books; monographs, dissertations and theses published in institutional repositories of renowned Brazilian and foreign universities. In addition, principles and concepts of public and private international law were analyzed, as well as the horizontal application of human rights and the terms of service of social media platforms. The result of the research suggests that the defense of freedom of expression has taken on more complex proportions, resulting in the need to find a content moderation arrangement that better shapes the practical aspects of online speech, with social media platforms having to adopt in their internal rules a content of fundamental rights, especially regarding freedom of expression, as well as evaluating scenarios with alternative solutions for content moderation, facing new standards of proportionality and probability to be explored, whose objective is to improve the process of content moderation.

KEYWORD: Fundamental Rights. Content Moderation. Social media. Alternative remedies.

LISTA DE ABREVIATURAS E SIGLAS

ARPA -*Advanced Research Projects Agency*

CEDH – Convenção Europeia de Direitos Humanos

CGI – Comitê Gestor da *Internet*

DCPR – Coalizão Dinâmica sobre a Responsabilidade das Plataformas

DMCA – *The Digital Millenium Copyright Act*

DNS – *Domain Name System*.

GAC – *Governmental Advisory Comitee*

IANA – *Internet Assigned Number Authority*.

ICANN – *Internet Corporation for Assigned Names and Numbers*.

IEEE – *Institute of Electrical and Electronics Engineers*.

IETF – *Internet Engineering Task Force*.

IGF – *Internet Governance Forum*.

IP – *Internet Protocol*.

ITU – *International Telecommunication Union*

NetzDG – *Netzwerkdurchsetzungsgesetz* - o *Network Enforcement Act*

ONU – Organização das Nações Unidas

PIDCP – Pacto Internacional sobre Direitos Civis e Políticos

RFC – *Request for Comments*

RIR – *Regional Internet Registries*

STF – Supremo Tribunal Federal

TSE – Tribunal Superior Eleitoral

W3C – *World Wide Web Consortium*.

WGIG – *Working Group on Internet Governance*.

WSIS – *World Summit on the Information Society*.

SUMÁRIO

INTRODUÇÃO	11
CAPÍTULO I – LIBERDADE DE EXPRESSÃO	21
1.1. Uma aproximação conceitual	21
1.2. Justificativas Constitucionais: liberdade de expressão como um direito fundamental	25
1.2.1. A busca da verdade e o Mercado livre de Ideias	27
1.2.2. Teoria da autossatisfação (<i>self-fulfilment</i>) individual.....	34
1.2.3. Teoria do autogoverno ou autodeterminação democrática.....	38
1.2.4. A liberdade de expressão como garantia da estabilidade governamental	42
1.2.5. Síntese conclusiva	44
1.3. Liberdade de expressão: um direito fundamental limitado?.....	44
1.4. A liberdade de expressão na era digital	53
1.4.1. A evolução do discurso <i>online</i>	54
1.4.2. Características do novo discurso <i>online</i>	57
CAPÍTULO II – GOVERNANÇA PRIVADA E LIBERDADE DE EXPRESSÃO NAS PLATAFORMAS DE MÍDIA SOCIAL	62
2.1. Governança (privada) da <i>internet</i> : definição.	63
2.2. Governança (privada) da <i>internet</i> : dimensão subjetiva	68
2.3. Governança (privada) da <i>internet</i> : dimensão objetiva.....	79
2.4. Governança das plataformas e direitos humanos: o tom constitucional dos termos de uso	82
2.4.1. Li e Aceito: termos de uso possuem tom constitucional	87
2.4.2. Termos de uso e a liberdade de expressão.....	93
2.4.3. O primeiro relatório sobre direitos humanos da <i>Meta</i> e a Declaração para o Futuro da <i>Internet</i> : avanços e perspectivas.	104
2.5. A eficácia horizontal dos direitos fundamentais na proteção da liberdade de expressão nas plataformas de mídia social.	110
CAPÍTULO III – MODERAÇÃO DE CONTEÚDO: para além da remoção	117
3.1. Moderação de conteúdo: principais tópicos	120
3.2. Moderação de conteúdo: indo além da remoção	133
3.3. Remédios alternativos de moderação	142
3.4. Aspectos normativos da moderação de conteúdo no Brasil	166
3.4.1. O Marco Civil da Internet.....	168
3.4.2. Proteção de Dados	174
3.4.3. Moderação de Conteúdo e Período Eleitoral	178
CONSIDERAÇÕES FINAIS.....	187
REFERÊNCIAS	197
ANEXO I	222

INTRODUÇÃO

A historiografia mais popular sobre o início da *internet* aponta a sua criação por militares norte-americanos do Departamento de Defesa dos Estados Unidos (ARPA - *Advanced Research Projects Agency*) os quais, com receio de terem segredos militares roubados por soviéticos, criaram um sistema de comunicação cujo compartilhamento de informações se dava entre atores geograficamente distantes. Posteriormente, foram incorporadas Universidades americanas ao projeto, propiciando, em outubro de 1969, a realização da primeira conexão entre pessoas dispostas em lugares diferentes. Após o desenvolvimento inicial da rede mundial de computadores (*rede*) nas décadas de 1980 e de 1990, surgiu uma nova fase de sua evolução, atingindo ampla divulgação perante a população, em decorrência de sua privatização, a qual iniciou-se por meio da transição do controle militar para o domínio civil, gerando novas possibilidades de comercialização. A sociedade da informação começava a se firmar no incremento das “supervias da informação”.¹

Nesse contexto, a *Word Wide Web* contribuiu para evolução e exerceu papel decisivo na concretização das promessas da *internet*, com o termo *web* passando quase a ser sinônimo de *internet*. Como argumenta Janet Abbate, a *web* completou a mutação da *internet*, a qual passou de ferramenta de pesquisa à mídia popular e, ao mesmo tempo, consolidou os seus princípios norteadores de descentralização, arquitetura aberta e participação ativa dos usuários, podendo cada indivíduo ser, simultaneamente, produtor e consumidor de informações.²

A partir dessa disseminação de elementos informativos, os *hackers* passaram a desempenhar um papel importante na história moderna do conceito de informação, em que pese uma visão negativa,³ muitas vezes, devido à falta de conhecimento da natureza técnica do objeto de sua dedicação. Dentre eles, existem pessoas conhecidas, como Bill Gates – na adolescência aprimorou sistemas de segurança computacional; os fundadores da *Apple*, Steve Jobs e Steve Wozniack – criaram sua empresa, devido ao *phreaking*, consistindo em desviar as linhas telefônicas e o criador do *Facebook*, Mark Zuckerberg – conhecido por manipular o banco de fotos dos estudantes da Universidade de *Harvard*, servindo de base para o

¹ LOVELUCK, Benjamin. **Redes, liberdades e controle:** uma genealogia política da *internet*. Petrópolis: Vozes, 2018. p. 114.

² ABBATE, Janet. **Inventing the Internet.** Cambridge: MIT Press, 1999. p. 217-218.

³ Loveluck explica que “apesar de refletir uma imagem de marginalidade, o hacking constitui a cultura comum aos atores principais da paisagem contemporânea da informática e da internet. Todos eles têm compartilhado determinados valores e determinadas práticas, pelo menos, em seus primeiros anos [...]” (LOVELUCK, Benjamin. **Redes, liberdades e controle:** uma genealogia política da *internet*. Petrópolis: Vozes, 2018. p. 128).

lançamento de sua empresa. Todos considerados *hackers* por desenvolverem percepções distintas da informática e sobretudo pela forma de se relacionar com a economia.⁴

O sucesso de suas determinações tornou-se um negócio bastante lucrativo, a partir da livre circulação da informação, combinado a um empreendimento rentável, único e verdadeiro, incentivando a inovação e permitindo a sua democratização. Com isso, a *internet* evoluiu, tornando-se um fator econômico e político, com cada vez mais relevância. Aos poucos, a promessa inicial de liberdade irrestrita, da rede, principal compromisso da *net*, passa a sofrer restrições.

É necessário enfatizar a quantidade surpreendente de dados gerados, obrigando aos programadores a criarem soluções para armazenamento dessas informações. A demanda por espaços digitais de armazenamento aumentou, intrincando suas tramas de tal forma que o ambiente digital⁵ hoje é um bem de valor de difícil mensuração, acarretando um novo potencial econômico, sendo considerado o novo petróleo.⁶

Esses enormes bancos de dados tornaram-se lucrativos também por permitirem mapear os *perfis* dos usuários endereçando-lhes publicidades e campanhas de *marketing*. Para tanto, os grandes atores da *web* incentivam manifestações dos usuários e suas interações, montando um perfil individualizado capaz de denotar suas preferências de compras e de serviços. Esse conhecimento engendra substanciais lucros, assumindo importância não só econômica, mas também é o acesso a esses dados que fomenta uma nova forma de poder.

Essa forma de poder foi demonstrada no conhecido escândalo envolvendo a subsidiária norte-americana do grupo *Strategic Communication Laboratories* (SCL), a *Cambridge Analytica*⁷, durante a eleição presidencial norte-americana, ou ainda na campanha

⁴ LOVELUCK, Benjamin. **Redes, liberdades e controle**: uma genealogia política da *internet*. Petrópolis: Vozes, 2018. p. 128.

⁵ “A explosão contínua de dados resultou em aproximadamente 40 trilhões de gigabytes (ou 40 zetabytes) somente em 2020. São 1,7 MB de dados criados a cada segundo para cada pessoa na Terra, aponta a IBM – International Business Machines Corporation”. (DAILEY, Doug. **Netezza and IBM Cloud Pak for Data: A knockout combo for tough data**. Disponível em: <https://www.ibm.com/blogs/journey-to-ai/2020/06/netezza-and-ibm-cloud-pak-a-knockout-combo-for-tough-data/>. Acesso em: 29/11/2021).

⁶ BRUZZONE, Andrés. **Ciberpopulismo**: política e democracia no mundo digital. São Paulo: Contexto, 2021. p. 45.

⁷ Para um estudo mais completo ver: SANTOS, Ivo Luís Azevedo da Costa. **How does innovation influence the relationship between the socio-technical change and policy mix: acknowledging lessons from the cambridge analytica case?** Dissertação (Mestrado). Faculdade de Economia. Universidade do Porto. Porto: 2019; OLIVEIRA, João António França de. **Rússia, Cambridge Analytica e as Eleições Presidenciais Norte-Americanas de 2016**: o ciberespaço como o mais novo domínio da conflitualidade política. Dissertação (Mestrado). Instituto de Estudos Políticos. Universidade Católica Portuguesa. Lisboa: 2020; BARCELOS, Júlia Rocha de. **Big Data, Algoritmos e Microdirecionamento**: desafios para a regulação da propaganda eleitoral. Dissertação (Mestrado). Universidade Federal de Minas Gerais, Faculdade de Direito, Belo Horizonte: 2019; e EDWARDS, L. *Cambridge Analytica and the deeper malaise in the persuasion industry*. Londres: **LSE Business Review**, 2018.

pelo *Brexit* – saída da Grã-Bretanha da Comunidade Europeia, utilizando, clandestinamente, dados de 87 milhões de perfis, indevidamente desviados do *Facebook*, com a finalidade de persuasão política, sem o conhecimento e o consentimento dos usuários⁸.

Constatou-se, na virada da década de 2000, o fato de a *internet* não ser um ambiente de liberdades ilimitada,⁹ mas sim uma arquitetura suscetível a mudanças, passando, então, a incorporar o debate sobre a necessidade de sua regulação, bem como a participação estatal nesse processo.

Nesse sentido, Balkin desenvolve a concepção de, no século XXI, a regulação do diálogo *online* ser pluralista, envolvendo vários atores emoldurados em um triângulo, dispondo-se da seguinte forma: o primeiro canto representando os Estados-nação, o segundo, as empresas privadas de infraestrutura de *internet* – dentre elas as empresas de mídia social, motores de buscas, provedores de *internet*, etc., em última parte, os usuários, organizações da sociedade civil e outros.¹⁰

Quanto ao regulamento do discurso, na perspectiva pluralista, Balkin classifica como “velha escola” e “nova escola”, nas quais as técnicas tradicionais ou antigas de regulação da fala tendem a utilizar penalidades criminais ou civis. Enquanto as novas técnicas regulam o controle sobre as redes digitais, alterando o foco de editores para proprietários de infraestrutura digital, apontando três características básicas da nova escola: censura colateral, cooperação ou cooptação pública/privada e restrição digital prévia.¹¹

⁸ Conferir a respeito: SHOWDEN, Edward. **Eterna Vigilância**: como montei e desvendei o maior sistema de espionagem do mundo. São Paulo: Planeta do Brasil, 2019. Ademais, sobre a aceitação dos “termos de uso” de uma rede, sem uma transparência adequada, o ex-Diretor de tecnologia da empresa *Cambridge Analytica*, Christopher Wylie, afirma em entrevista em 2018 ao El País: “Não há uma dicotomia entre público e privado, é um espectro. O problema é que se você diz que só é privada a comunicação de um com um outro, e tudo o mais é livre para todos, se estendermos essa lógica poderiam pôr câmeras de vigilância na sala da sua casa. Isso é o equivalente digital da sua sala. As pessoas se relacionam no Facebook como se falassem com amigos. Não estão publicando para que o mundo veja. O fato de que você não vá comprovar as condições padronizadas de privacidade não é motivo para não respeitar essa expectativa razoável de que há certo grau de privacidade no que você publica.” (WYLIE, Christopher, *apud* GUIMÓN, Pablo. “O ‘Brexit’ não teria acontecido sem a Cambridge Analytica”. Disponível em: https://brasil.elpais.com/brasil/2018/03/26/internacional/1522058765_703094.html. Acesso em: 30/11/2021).

⁹ A primeira foi o estouro da bolha da *internet* que assistiu ao colapso da Nova Economia, prometia um crescimento infinito, no modelo multiplicidade de microempresas, contudo somente algumas sobreviveram, num modelo concentrado. O segundo evento foi uma série de processos e intervenções judiciais, exemplo emblemático é processo contra o Yahoo! Na França, por discursos racistas, antisemitas ou xenófobos. O terceiro evento foi os atentados de 11 de setembro de 2001, com a vigilância cada vez mais estrita da rede tanto por empresas privadas quanto por agências de inteligência. (LOVELUCK, Benjamin. **Redes, liberdades e controle**: uma genealogia política da *internet*. Petrópolis: Vozes, 2018. p. 165-167).

¹⁰ BALKIN, Jack M. Free Speech is a triangle. **Columbia Law Review**. v. 118, p. 2011–2055, 2018. p. 2011. Disponível em: <https://openyls.law.yale.edu/bitstream/handle/20.500.13051/17953/Balkin%2C%20Free%20Speech%20is%20a%20Triangle.pdf?sequence=1&isAllowed=y>. Acesso em: 04/11/2021.

¹¹ BALKIN, Jack M. Old-School/New-School Speech Regulation. **Harvard Law Review**. v. 127, n. 8, p. 2296-2342, 2014. p. 2298-2299. Disponível em:

Nessa configuração triangular, Balkin aponta três principais pontos problemáticos, (i) os Estados-nação tendem a pressionar as empresas de infraestrutura *online*, por meio da nova escola de regulação do discurso, criando problemas de censura colateral e de restrição digital prévia; (ii) as empresas de mídia social desenvolveram governança privada complexa, direcionando os usuários finais, sem observar o devido processo e sem transparência; (iii) os usuários das plataformas ficam vulneráveis, tanto à vigilância quanto à manipulação digital.¹²

Destarte, as plataformas digitais exsurtem como mais um ator importante na sociedade, imprimindo inevitável impacto no cotidiano social, além de moldar novos hábitos e gerar necessidades de padrão desmaterializadas. Destacando-se, no contexto virtual, como ente da esfera de característica pública,¹³ dentro de um ambiente(negócio) privado,¹⁴. Esse paradoxo, o qual, além de inovador e benéfico, pode tornar-se ofensivo aos valores democráticos e aos direitos fundamentais.

Nesse contexto, surge um dos principais debates sobre as plataformas da *internet*, o poder de decisão delas e a influência da moderação de conteúdo no discurso *online*. Por conta desse poder exercido pelas plataformas *online*, Kate Klonick denominou-os como “Os Novos Governadores”.¹⁵ Nessa linha, Mark Zuckerberg afirma que o *Facebook* “é mais como um governo do que uma empresa tradicional. Temos essa grande comunidade de pessoas e, mais do que outras empresas de tecnologia, estamos realmente definindo políticas.”¹⁶

https://openyls.law.yale.edu/bitstream/handle/20.500.13051/4406/127_Harv_L_Rev._2296_2014.pdf?sequence=2&isAllowed=y. Acesso em: 04/11/2021.

¹² BALKIN, Jack M. Free Speech is a triangle. **Columbia Law Review**. v. 118, p. 2011–2055, 2018. p. 2011. Disponível em: <https://openyls.law.yale.edu/bitstream/handle/20.500.13051/17953/Balkin%2C%20Free%20Speech%20is%20a%20Triangle.pdf?sequence=1&isAllowed=y>. Acesso em: 04/11/2021.

¹³ Para Balkin “as redes de comunicação são públicas por natureza, mesmo que sua infraestrutura tecnológica seja de propriedade privada. Portanto, eles devem conceder acesso justo às suas redes, não devem agir como gargalos e não devem discriminar injustamente o conteúdo de outras fontes.” Tradução livre do original: (BALKIN, Jack M., Digital Speech and Democratic Culture: A Theory of Freedom of Expression for the Information Society, **New York University Law Review**, v. 79, p. 1–55, 2004. p. 50. Disponível em: https://openyls.law.yale.edu/bitstream/handle/20.500.13051/1712/Digital_Speech_and_Democratic_Culture_A_Theory_of_Freedom_of_Expression.pdf?sequence=2. Acesso em 04/11/2021).

¹⁴ Nesse sentido, Gustavo Binenbojm: “As redes sociais não são espaços reservados comparáveis aos cômodos de uma residência, às salas de um escritório privado ou aos arquivos de computadores e smartphones particulares. A despeito de sua natureza não estatal, o Facebook é um fórum público que envolve uma pluralidade de pessoas em convívio livre, a partir da decisão autônoma de cada um de solicitar o acesso e aceitar a companhia alheia.” BINENBOJM, Gustavo. **Liberdade igual: o que é e por que importa**. 1 ed. Rio de Janeiro: História Real: 2020. p. 65.

¹⁵ Conferir: KLONICK, Kate. The New Governors: The People, Rules and Processes Governing Online Speech. **Harvard Law Review**. v. 131. p. 1598-1670. 2018. p. 1603. Disponível em: https://harvardlawreview.org/wp-content/uploads/2018/04/1598-1670_Online.pdf. Acesso em: 28/05/2022.

¹⁶ KIRKPATRICK, David. The Facebook Effect: The Inside Story of the Company That is Connecting the World. New York: Simon & Schuster. 2010. **Zuckerberg Transcripts**. 67. Disponível em: https://epublications.marquette.edu/zuckerberg_files_transcripts/67. Acesso em: 18/05/2022.

As plataformas *online* representam agentes que integram a governança da *Internet*, sobretudo por meio do exercício da “moderação de conteúdo”, constituindo regras e arranjos utilizados pelas plataformas para remover, limitar alcance, rotular conteúdo como desinformação, suspender ou remover conteúdo ou contas, dentre outros.

A moderação de conteúdo desafia uma série de problemas e debates, como por exemplo, remoções de conteúdo cujo teor não viola as regras da plataforma (falso positivo), a manutenção de conteúdo contrário às regras (falso negativo); responsabilidade algorítmica; legitimidade; neutralidade da rede; moderação automatizada; saliento que a ausência de moderação pode causar prejuízos aos usuários ou silenciá-los. Além disso, as plataformas têm sido cobradas para manter um ambiente digital sadio e, do mesmo modo, criticadas por possuírem uma postura interferente ou cerceadora da liberdade de expressão de seus usuários.

Nesse estudo, foi realizado um corte metodológico, concentrando-se no exame dos atores privados (ocupantes do *segundo canto do triângulo* de Balkin), especificamente nas plataformas *online*, quanto à governança privada criada por essas empresas (*segundo ponto problemático* apontando por Balkin). Ressalte-se que, ao referir-se às plataformas, não há uma especificidade, fala-se de maneira geral, pois as tratativas ou remédios analisados durante a pesquisa podem ser aplicados a todos os problemas encontrados, de acordo com o caso e o seu formato/modelo.

Note-se ainda, especialmente em relação à sua atuação na *moderação de conteúdo* ou à tomada de decisão para correção de determinada ação, sem importar se é de um usuário ou conta, se violarem as regras internas das plataformas, os *remédios alternativos* serão apresentados, como forma de incentivar os moderadores a adotarem medidas tratativas, *além da simples remoção* do conteúdo.

Nesse panorama, o interesse em analisar a moderação de conteúdo, embora comportando a remoção de conteúdo *online* como sua principal ferramenta, despertou em decorrência de eventos nos quais as plataformas de redes sociais removem conteúdos e/ou contas de usuários por violação de seus termos de uso, serviço ou políticas, sem, contudo, analisar o contexto. A maior observação a respeito é que tais retiradas ocorrem sem ouvir o usuário, sem um devido processo previamente estabelecido entre as partes, e, sobretudo, sem ponderar a possibilidade de adotar ferramentas alternativas, cujo conteúdo/conta não sejam removidos permanentemente. Essa abordagem pode ferir a liberdade de expressão, além de outros direitos fundamentais.

O evento mais notório sobre a regulamentação da plataforma de mídia social, ocorreu em 08 de janeiro de 2021, quando o então Presidente dos Estados Unidos da América, Donald

Trump, sofreu uma sanção de suspensão permanente da sua conta no *Twitter*. Posteriormente, a plataforma publicou um texto no qual explicava aos usuários os motivos da punição definitiva de Trump e, embora os dois *tweets* fossem relativamente banais, separadamente, ao serem contextualizados, demonstraram fazer parte de eventos mais amplos naquele país, em especial, o episódio da invasão do Capitólio Americano, dois dias antes da suspensão da conta de Trump, entretanto, os *tweets* estavam sendo interpretados dentro e fora daquela rede social, não se restringindo ao ambiente do *Twitter*, servindo como um incentivo a mais violência.¹⁷

O *Facebook*, usando o mesmo raciocínio, também suspendeu indefinidamente as contas de Trump em suas plataformas; o CEO Mark Zuckerberg manifestou-se na própria rede social por meio de um *post* público, no qual anunciou a decisão.¹⁸ Contudo, duas semanas depois, encaminhou essa decisão ao seu Conselho de Supervisão.¹⁹

Algumas das fontes de embasamento deste trabalho surgiram a partir de eventos ocorridos no Brasil, cuja importância histórica se deu às vésperas das eleições gerais de 2018, quando o *Facebook* passou por uma “rigorosa investigação”, resultando na desativação de, aproximadamente, 196 páginas e 87 contas de usuários vinculados ao Movimento Brasil Livre (MBL),²⁰ com a fundamentação de utilização de perfis para “gerar divisão e espalhar desinformação”, violando a política de autenticidade da rede social.

Noutro episódio, em outubro de 2021, o *Facebook*, *Instagram* e *YouTube* removeram uma *live* do Presidente da República, Jair Messias Bolsonaro, sob a alegação de o presidente ter infringido a política das plataformas, com a informação de, durante a transmissão, associar, falsamente, a vacinação contra covid-19 à AIDS.²¹

Recentemente, em 10 de agosto de 2022, o *Youtube* removeu de sua plataforma uma *live*, do evento realizado em 18 de julho de 2022, no qual o presidente Jair Bolsonaro, reuniu-se com embaixadores no Palácio do Planalto. Durante o evento, o chefe do executivo brasileiro fez ataques às urnas eletrônicas, colocando em dúvida o processo eleitoral brasileiro, sem apresentar provas, além disso, fez críticas ao Supremo Tribunal Federal e ao

¹⁷ TWITTER INC. Permanent suspension of @realDonaldTrump. **Blog.twitter**. 2021. Disponível em: https://blog.twitter.com/en_us/topics/company/2020/suspension. Acesso em: 04/11/2021.

¹⁸ ZUCKERBERG, Mark. Post Announcing Indefinite Suspension of Trump's Accounts. **Facebook**. Disponível em: <https://www.facebook.com/zuck/posts/10112681480907401>. Acesso em: 06/11/2021.

¹⁹ O caso é tratado no Capítulo III.

²⁰ GONÇALVES, Gêssica Brandino; RESENDE, Sarah Mota. Facebook retira rede de páginas e perfis do ar e atinge MBL. 25 de julho de 2018. **Folha de São Paulo**. Disponível em: <https://www1.folha.uol.com.br/poder/2018/07/facebook-retira-rede-de-paginas-e-perfis-do-ar-e-atinge-mbl.shtml>. Acesso em: 29/11/2021.

²¹ MIRANDA, Cássia. Facebook e Instagram excluem live de Bolsonaro com fake news sobre aids e a vacina da covid. **O Estado de São Paulo**. Disponível em: <https://politica.estadao.com.br/noticias/geral,facebook-e-instagram-excluem-live-de-bolsonaro-com-fake-news-sobre-aids-e-a-vacina-da-covid,70003879214>. Acesso em 29/11/2021.

Tribunal Superior Eleitoral. O *Youtube* informou que a remoção ocorreu devido à atualização nas diretrizes de combate à desinformação sobre processo eleitoral no *YouTube*.²²

Acrescenta-se, por fim, um exemplo brasileiro de falha na moderação, quando, em outubro de 2020, houve a remoção do conteúdo de um usuário, o qual publicou oito fotografias mostrando sintomas de câncer de mama, sendo que das imagens exibidas, em cinco apareciam mamilos e a legenda da postagem era sobre o apoio à campanha Outubro Rosa (Campanha de prevenção ao câncer de mama). Contudo, apesar de as Políticas sobre Nudez do *Facebook* incluir a exceção de publicações nas quais haja mamilos, desde que o propósito seja o de conscientizar o público sobre o câncer de mama, o *post* foi removido pela moderação automatizada, aplicando o Padrão da Comunidade sobre Nudez Adulta e Atividade Sexual do *Facebook*. Devido à falha, o caso foi encaminhado para o Comitê de Supervisão do *Facebook*, e, embora o próprio *Facebook*, tenha reconhecido o erro, mesmo antes da decisão colegiada, o Comitê reforçou o risco que a moderação automatizada pode trazer à liberdade de expressão.²³

Com efeito, a importância deste trabalho é contribuir para o debate da restrição de conteúdo, com utilização do recurso de remoção de conteúdo e/ou perfis/contas de usuários, sem prévia oitiva, motivado tão somente por regras internas da plataforma digital. Sendo pertinente a reflexão sobre tal conduta não configurar, de fato, uma superposição de um ator privado, gerenciando a vida (política) de uma sociedade. Questionando-se se a remoção de conteúdo das redes sociais, seguindo seus próprios termos de uso, seria o único instrumento condizente sob a ótica dos direitos fundamentais e, se assim fosse, afigurar-se-ia que somente a prática binária: remover ou não um conteúdo seria suficiente para a prática da moderação, torna-se essencial, portanto, o aprofundamento desse pensamento para além da simples remoção, indicando-se uma outra via diferente, examinando um catálogo de remédios alternativos, como forma de atenuar o “interesse privado”, quando colidirem com direitos fundamentais, dentre eles: a liberdade de expressão.

Esta pesquisa situa-se no debate da governança da *Internet*, sob a perspectiva da regulação do discurso pelas plataformas digitais, por meio de seus termos de uso e da governança privada de moderação de conteúdo. Nesse contexto, os seguintes problemas são postos em destaque na pesquisa, cuja força motriz motivou e construiu a base, a argumentação e o norteamento deste trabalho: *É possível as redes sociais removerem conteúdos de*

²² NIKLAS, Jan. YouTube remove livre de Bolsonaro com embaixadores. **Yahoo.com**. 2022. Disponível em: <https://esportes.yahoo.com/youtube-remove-live-bolsonaro-com-210850760.html>. Acesso em 11/08/2022.

²³ CONSELHO DE SUPERVISÃO. Decisão do sobre o caso 2020-004-IG-UA. **Oversightboard.com**. 2021. Disponível em: <https://www.oversightboard.com/decision/IG-7THR3SI1>. Acesso em: 10/05/2022.

usuários? Sendo positivo, questiona-se também se é o único meio hábil para corrigir violações às regras das plataformas, garantindo o uso racional da liberdade de expressão nas redes sociais?

Ressalta-se o objetivo geral ser a *compreensão das principais controvérsias sobre a operação da governança privada na moderação de conteúdo e como o direito brasileiro trata tal problemática.*

Nessa perspectiva, para atingir as respostas para os problemas da pesquisa, os objetivos são: *fornecer* um panorama da liberdade de expressão no ambiente virtual, identificando novas ameaças à higidez do direito fundamental, bem como uma nova postura do Estado como garantidor deste direito fundamental. *Compreender* como opera a governança privada das plataformas digitais, analisando os atores particulares; sobretudo quanto às regras elaboradas pelas plataformas, aceitas pelos usuários em contrato de adesão, em concordância ou em conflito com os direitos fundamentais; bem como, *identificar* os principais tópicos de discussão sobre a moderação de conteúdo nas redes sociais, apresentando o rol de remédios para correção de conteúdos ou de ações que violem o regramento interno das empresas de mídia social, para além da simples remoção de conteúdo.

Metodologicamente, este estudo realiza uma minuciosa pesquisa bibliográfica e documental, cuja revisão e análise da literatura existentes não só fomentam os principais temas desta dissertação, como também propõe uma reflexão sobre os problemas encontrados na moderação de conteúdo, buscando sugerir soluções ou melhorias, principalmente, no que tange à liberdade de expressão, governança da *Internet*, moderação de conteúdo e mídia social. Embora reconheça-se a importância da literatura sobre governança da *Internet*, do ponto de vista dos atores privados, como novos governadores do discurso *online*, parece plausível que a observação dos métodos de moderação particular coexista com a melhoria contínua do seu próprio arcabouço teórico. Saliente-se que os moderadores ou os responsáveis pelas plataformas realizam a elaboração de normas de uso, a fim de tentar prever os casos em que o exercício da moderação de conteúdo e a incidência de direitos fundamentais, sobretudo a liberdade de expressão, sejam preservados. A revisão da literatura realizada inclui também a da legislação; termos de uso e serviço de mídia social; declarações de direitos digitais e outros instrumentos internacionais; artigos científicos disponíveis em periódicos nacionais e estrangeiros; livros; monografias, dissertações e teses publicadas em repositórios institucionais de renomadas universidades brasileiras e de outros países.

Durante a fase de pré-pesquisa, foram definidas as palavras-chave e selecionadas as bases de dados, a fim de encontrar a literatura mais relevante, pois o tema pesquisado está

situado numa área do direito de natureza interdisciplinar, incorporando trabalhos de constitucionalistas, teóricos do direito, economistas, profissionais digitais e jornalistas.

A pesquisa foi de natureza qualitativa. Além disso, foram analisados princípios e conceitos de direito internacional público e direito internacional privado, bem como a aplicação horizontal dos direitos humanos e dos termos de serviço das plataformas de mídia social.

O referencial teórico é ancorado nos estudos de Jack M. Balkin (*Yale Law School*) sobre liberdade de expressão como meio de participação e interação na formação de uma cultura democrática, na qual, conforme a visão de Balkin, os indivíduos têm uma oportunidade justa de participar das formas de construção de significado que os constituem como indivíduos. Além da nova visão sobre a liberdade de expressão na era digital e uma nova forma como essa liberdade pode ser protegida, segundo Balkin, por meio de infraestrutura tecnológica e regulatória.²⁴ Bem como, nos ensinamentos de Eric Goldmen (*Santa Clara University School of Law*), sobretudo em relação à moderação de conteúdo como meio de governança privada das plataformas *online*, e no estudo de remédios alternativos na moderação.

O trabalho é dividido em três capítulos, além de introdução e considerações finais. O primeiro capítulo analisa os principais aspectos da liberdade de expressão, apresentando uma nova visão oriunda do avanço tecnológico, apontando os contornos necessários para privilegiar os valores do discurso, perante a utilização em massa de redes sociais. Com essa nova realidade, proporcionando uma dualidade de maior prestígio da funcionalidade da mídia social com a liberdade de expressão e, ao mesmo tempo, a falta ou má condução das plataformas digitais em assegurar os valores da liberdade de expressão, em detrimento de seus meios próprios de gerenciar a mídia social.

²⁴ “A era digital torna cada vez mais evidente o que sempre foi assim – que o sistema de liberdade de expressão depende de algo mais do que a soma de todos os direitos individuais de liberdade de expressão. Conta com infraestrutura tecnológica e regulatória. Essa infraestrutura é produzida por meio de regulamentação governamental, por meio de subsídios e programas de direitos e por meio de projetos tecnológicos. A liberdade de expressão é, e sempre foi, um fenômeno cultural, bem como jurídico ou político.” Tradução livre do original: “The digital age makes increasingly apparent what has Always been the case — that the system of free expression relies on something more than the sum of all individual free speech rights. It relies on a technological and regulatory infrastructure. That infrastructure is produced through government regulation, through government subsidies and entitlement programs, and through technological design. Freedom of speech is, and always has been, a cultural phenomenon as well as a legal or political one. (BALKIN, Jack M., *Digital Speech and Democratic Culture: A Theory of Freedom of Expression for the Information Society*, **New York University Law Review**, v. 79, p. 1–55, 2004. p. 48. Disponível em: https://openyls.law.yale.edu/bitstream/handle/20.500.13051/1712/Digital_Speech_and_Democratic_Culture_A_Theory_of_Freedom_of_Expression.pdf?sequence=2. Acesso em 04/11/2021).

No segundo capítulo, identifica-se como as empresas de mídia social equilibram considerações de direitos fundamentais em suas plataformas, particularmente em relação à liberdade de expressão, trazendo conceitos de governança da *internet*, dentro desta, abordando a governança pelas plataformas, analisando os termos de uso e serviço, verificando a constitucionalidade em sua elaboração e se há promoção à proteção da liberdade de expressão. Finalizando o capítulo com uma abordagem sobre a eficácia horizontal dos direitos fundamentais na proteção da liberdade de expressão nas plataformas de mídia social.

No andamento, o terceiro capítulo apresenta argumentos para utilização de remédios ou soluções alternativas na correção de conteúdo, ou as ações *online* que violem as regras internas, apresentando ferramentas alternativas para a moderação de conteúdo, para além da remoção, como forma de prestigiar a liberdade de expressão. Por fim, realiza-se uma análise sucinta sobre a normativa brasileira quanto a moderação de conteúdo, apontando rumos para aplicação de soluções alternativas além da remoção de conteúdo.

CAPÍTULO I – LIBERDADE DE EXPRESSÃO

A análise da liberdade de expressão – fulcro desse capítulo e matéria complexa, sendo objeto de estudos recorrente, dada sua importância no desenvolvimento de uma sociedade democrática, que tem sofrido graves convulsões, sobretudo na *ágora* das mídias sociais – torna-se pertinente, traçando linhas gerais sobre sua conceituação, abordando concepções filosóficas e delimitações teóricas para melhor compreensão do tema.

Cumprindo, preliminarmente, esclarecer que será utilizado no trabalho o termo genérico liberdade de expressão, conforme consagrado na Declaração Universal dos Direitos Humanos. Com isso, a pesquisa se funda na perspectiva de livre manifestação do pensamento, opiniões e informações, além da liberdade de comunicação social. Não sendo objeto de atenção, liberdade de ensino e pesquisa, liberdade de reunião e manifestação, liberdade artística e científica, além da liberdade religiosa.

1.1. Uma aproximação conceitual

Os Estados Democráticos priorizam dentro de seu arcabouço jurídico a liberdade de expressão (e de informação) como um dos pilares de sua sustentação, buscando assegurar sua plena observância como meio de auto complementar o princípio democrático e, com isso, permitindo a ambos se manterem reciprocamente. Michelman explica a democracia e a liberdade de expressão possuindo relação recíproca condicionante e assumindo um caráter complementar, dialético e dinâmico, na medida em que, ora mais liberdade de expressão signifique mais democracia e vice-versa; de maneira que a liberdade de expressão pode desencadear riscos a democracia e, de igual forma, esta pode restringir a liberdade de expressão.²⁵

Nessa perspectiva, há um consenso acerca da imprescindibilidade da liberdade de expressão como uma das formas de Liberdades, figurando no catálogo de direitos fundamentais. De fato, é mais do que um direito, como aduz Torres, se referindo a liberdade de expressão como podendo “ser entendida como um conjunto de direitos relacionados às liberdades de comunicação.”²⁶ Assim, dentre as várias formas de expressão, o direito de

²⁵ MICHELMAN, Frank. Relações entre democracia e liberdade de expressão: discussão de alguns argumentos. In: SARLET, Ingo Wolfgang (org.). **Direitos fundamentais, informática e comunicação**. Porto Alegre: Livraria do Advogado, 2007. p. 49 e seguintes.

²⁶ TÔRRES, Fernando Carolina. O direito fundamental à liberdade de expressão e sua extensão. **Revista de Informação Legislativa**. Ano 50, número 200, out/dez, 2013, p. 62.

liberdade de expressão pode ser compreendido como a união de diversas liberdades fundamentais visando à “proteção daqueles que emitem e recebem informações, críticas e opiniões”.²⁷ Desta forma, há o entendimento de a liberdade de expressão, na ordem jurídica contemporânea, consistir, em sentido amplo, “num conjunto de direitos relacionados às liberdades de comunicação, que compreende: a liberdade de expressão em sentido estrito, a liberdade de criação e de imprensa, bem como o direito de informação.”²⁸

O constitucionalismo brasileiro, tradicionalmente, aderiu não apenas um elenco de direitos de liberdade específico, como também a liberdade de expressão, de reunião e manifestação, entre outras. Tendo adotado também, como afirmam Sarlet, Mitidiero e Marinoni, “um direito geral de liberdade, que atua como critério material para a identificação de outras posições jurídicas fundamentais, em especial como parâmetro para a dedução de liberdades específicas que não foram objeto de direito e expressa previsão pelo constituinte.”²⁹

Desde 2012, a Organização das Nações Unidas (ONU) tem reportado que os direitos humanos, incluindo a liberdade de expressão, devem ser protegidos tanto de forma *online*, como também *offline*.³⁰ Para determinar a observância pelos Estados, o artigo 19 do Pacto Internacional sobre Direitos Civis e Políticos (PIDCP)³¹ prevê a obrigatoriedade de os estados garantirem um ambiente propício para a liberdade de expressão além de proteger o seu exercício (parágrafo 1).³² A liberdade de expressão não é absoluta, mas qualquer restrição deve atender aos critérios: “previstos em lei e necessários” para a proteção dos direitos ou reputações individuais, da ordem pública ou da saúde e da moral públicas (parágrafo 3).³³ De acordo com o artigo 20 do PIDCP, uma restrição legítima à liberdade de expressão é “ódio nacional, racial ou religioso que constitui incitamento à discriminação, hostilidade ou

²⁷ TÔRRES, Fernando Carolina. O direito fundamental à liberdade de expressão e sua extensão. **Revista de Informação Legislativa**. Ano 50, número 200, out/dez, 2013, p. 62

²⁸ TÔRRES, Fernando Carolina. O direito fundamental à liberdade de expressão e sua extensão. **Revista de Informação Legislativa**. Ano 50, número 200, out/dez, 2013, p. 62.

²⁹ SARLET, Ingo Wolfgang; MITIDIERO, Daniel; MARINONI, Luiz Guilherme. **Curso de direito constitucional**. 9 ed. São Paulo: Saraíva Educação, 2020. p. 503.

³⁰ Conselho de Direitos Humanos das Nações Unidas (UNHRC). (2012, 16 de julho). **A promoção, proteção e prazer dos direitos humanos na internet** (A/HRC/20/8). Disponível em: <https://digitallibrary.un.org/record/731540?ln=en>. Acesso em 30/10/2021.

³¹ ICCPR (1966, 16 de dezembro). **Pacto internacional sobre direitos civis e políticos**. United Nations Human Rights Office of the High Commissioner.. Disponível em: <https://www.ohchr.org/en/professionalinterest/pages/ccpr.aspx>. Acesso em: 30/10/2021.

³² Texto livre do original: “Article 19: 1. Everyone shall have the right to hold opinions without interference.”

³³ Texto livre do original: “Article 19: 3. The exercise of the rights provided for in paragraph 2 of this article carries with it special duties and responsibilities. It may therefore be subject to certain restrictions, but these shall only be such as are provided by law and are necessary: (a) For respect of the rights or reputations of others; (b) For the protection of national security or of public order (ordre public), or of public health or morals.”

violência” (parágrafo 2).³⁴ Da mesma forma, o Pacto obriga os estados a implementarem e fazerem cumprir medidas apropriadas e eficazes para prevenir e proteger contra atos de discriminação por vários motivos, incluindo sexo, raça, cor, descendência ou origem nacional (artigo 2, parágrafo 1).³⁵

Anote-se, ainda, que outros diplomas normativos têm previsão expressa sobre o direito à liberdade de expressão, como se depreende do Artigo XIX da Declaração Universal dos Direitos do Homem, de 1948,³⁶ além do Artigo 13.1 da Convenção Americana de Direitos Humanos, conhecido como Pacto de São José da Costa Rica.³⁷

No Brasil, a liberdade de expressão sempre teve previsão nas constituições brasileiras, desde a Carta Imperial de 1824 (Art. 179, IV), bem como na Constituição de 1891 (Art. 72, § 12), na Constituição de 1934 (Art.113), na Constituição de 1937 (Art. 122), na Constituição de 1946 (Art. 141, § 5º), na Constituição de 1967 (Art. 150, § 8º), através da EC n. 1/1969 (Art. 153, § 8º). Nesse contexto, Robl Filho e Sarlet lecionam que com a Constituição Federal de 1988:

[...] as liberdades de expressão (ou liberdades comunicativas) foram objeto de mais detalhada positivação, passando a corresponder, pelo menos de acordo com texto constitucional, ao patamar de reconhecimento e de proteção compatível com um autêntico Estado Democrático de Direito.³⁸

Em linhas gerais, os norte-americanos trabalham com a “liberdade de fala” (*free speech*). Os germânicos adotam a “liberdade de comunicação” (*Kommunikationsfreiheit*). Enquanto a Declaração Universal dos Direitos Humanos consagra a “liberdade de pensamento e expressão”. O sistema europeu de proteção dos direitos humanos utiliza apenas o termo

³⁴ Texto livre do original: “Article 20: 1. Any propaganda for war shall be prohibited by law. 2. Any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence shall be prohibited by law.”

³⁵ Texto livre do original: “Article 2: 1. Each State Party to the present Covenant undertakes to respect and to ensure to all individuals within its territory and subject to its jurisdiction the rights recognized in the present Covenant, without distinction of any kind, such as race, colour, sex, language, religion, political or other opinion, national or social origin, property, birth or other status.”

³⁶ ONU. Declaração Universal dos Direitos do Homem, de 1948. “Artigo XIX. Toda pessoa tem direito à liberdade de opinião e de expressão; este direito inclui a liberdade de, sem interferência, ter opiniões e de procurar, receber e transmitir informações e idéias por quaisquer meios e independentemente de fronteiras.”

³⁷ OEA. Convenção Americana de Direitos Humanos. Pacto de São José da Costa Rica, de 1969. “Artigo 13.1. Toda pessoa tem direito à liberdade de pensamento e de expressão. Esse direito compreende a liberdade de buscar, receber e difundir informações e idéias 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.”

³⁸ ROBL FILHO, Ilton Norberto; SARLET, Ingo W. Estado democrático de direito e os limites da liberdade de expressão na Constituição Federal de 1988, com destaque para o problema da sua colisão com outros direitos fundamentais, em especial, com os direitos de personalidade. **Constituição, Economia e Desenvolvimento: Revista da Academia Brasileira de Direito Constitucional**. v. 8, n. 14, p. 112-142, Curitiba, 2016, p. 118.

“liberdade de expressão”. O sistema interamericano de direitos humanos segue a mesma senda da Declaração Universal.³⁹

Na atual Constituição Federal brasileira, a liberdade de expressão é abordada em sentido amplo, abrangendo a liberdade de “manifestação do pensamento”, a liberdade para exercício de “atividade intelectual, artística, científica e de comunicação”, dedica, ainda, um capítulo para a “comunicação social”, sendo este separado da seção dedicada à proteção dos direitos e garantias fundamentais. Incluindo, na liberdade de expressão, até mesmo, o direito fundamental às liberdades de consciência e de crença.⁴⁰

Nesse panorama, compreendemos a configuração de um aglomerado de direitos comunicativos fundamentais, dentre eles: a liberdade de expressão estrito senso e a liberdade de informação e de comunicação individual.⁴¹ Tal distinção apresenta importância prática para delimitar meios de proteção e de limitação desses direitos. A garantia da liberdade de expressão não está vinculada necessariamente à veracidade da opinião, somente há o dever de prová-la quando se trata da liberdade de informação, posto que essa se refere a fatos, desta forma para a sua veiculação é necessária a realização de checagem prévia de veracidade.⁴² Como visto, a liberdade de informação tem a verdade como seu limite, seja interno ou externo, pois sua finalidade é repassar ao ouvinte uma realidade fática.⁴³

Conceitualmente, a liberdade de expressão compreende tanto o direito (faculdade) de a pessoa se exprimir, de informar e de ser informado, quanto o de não se expressar ou mesmo de não se informar.⁴⁴

Como delineado, não há uma linha uniforme de conceituação de liberdade de expressão, variando conforme a concepção teórica adotada. Para melhor compreensão das raízes que sustentam esse princípio é salutar, a devida contextualização de suas bases históricas e filosóficas.

³⁹ FERREIRA, Felipe Grizoto. **Liberdade de expressão na era digital: desafios, perspectivas e aplicações**. Dissertação (Mestrado em Direito). Campinas: PUC-Campinas, 2021.

⁴⁰ Nesse sentido verificar: FACHIN, Luiz Edson. **Direito fundamental e expressão religiosa: entre a liberdade, o preconceito e a sanção**. In: MENEZES, Joyceane Bezerra de; TEPEDINO, Gustavo (Coord.). *Autonomia privada, liberdade existencial e direitos fundamentais*. Belo Horizonte: Fórum, 2019.

⁴¹ CANOTTILHO, José Joaquim Gomes; MACHADO, Jónatas E.M; GAIO JÚNIOR, Antônio Pereira. **Biografia não autorizada versus liberdade de expressão**. 3. ed. Curitiba: Juruá, 2017, p. 26.

⁴² CHEQUER, Cláudio. **A liberdade de expressão como direito fundamental preferencial prima facie**. Rio de Janeiro: Lumen Juris, 2011, p. 12.

⁴³ SIMÃO, José Luiz de Almeida; RADOVALHO, Thiago. A fundamentalidade do direito à liberdade de expressão: as justificativas instrumental e constitutiva para a inclusão no catálogo dos direitos e garantias fundamentais na constituição federal de 1988. **Cadernos do Programa de Pós-Graduação em Direito PPGD/UFRGS**. Edição Digital. v. 12, n. 1. 2017, p. 203-229, p.211.

⁴⁴ BRANCO, Paulo Gustavo Gonet. **Liberdades**. In: MENDES, Gilmar Ferreira; BRANCO, Paulo Gustavo Gonet. **Curso de direito constitucional**. 15. ed., São Paulo: Saraiva, 2020. p. 269.

No tópico seguinte, apresenta os fundamentos da liberdade de expressão como um direito fundamental, os quais, de igual modo, servem para demonstrar as argumentações das finalidades/necessidades atribuídas a liberdade de expressão.

1.2. Justificativas Constitucionais: liberdade de expressão como um direito fundamental

As Nações democráticas entendem a liberdade de expressão como algo imprescindível para a sociedade, figurando como uma das mais importantes das Liberdades. Nesse sentido, Sarlet afirma “é amplamente reconhecido que a liberdade de expressão constitui um dos direitos fundamentais mais preciosos e integra os catálogos constitucionais desde a primeira fase do constitucionalismo moderno.”⁴⁵

Existem alguns fundamentos que justificam a liberdade de expressão em sentido amplo como um direito fundamental, inserido, de alguma maneira, em duas grandes ordens, conforme deduz Dworkin,⁴⁶ na primeira inserção, a liberdade de expressão possui importância meramente instrumental, constituindo-se um meio pelo qual atinge um fim importante; enquanto na segunda categoria a liberdade de expressão, reveste-se de importância *per si*, não necessitando das consequências que esse direito pode desenvolver, seu desiderato como um fim em si mesmo.⁴⁷

Nesse passo, importante assinalar os quatro pontos utilizados pela doutrina para fundamentar a liberdade de expressão como um direito fundamental, sistematizados por Barendt, a saber:

- a) a liberdade de expressão é um direito fundamental em razão de ser um importante instrumento para a descoberta da verdade; b) a liberdade de expressão é um direito fundamental em razão de garantir uma autossatisfação (*self-fulfilment*) individual; c)

⁴⁵ SARLET, Ingo Wolfgang; WEINGARTNER NETO, Jayme. Liberdade de expressão: algumas ponderações em matéria penal à luz da Constituição Federal do Brasil. **Revista Espaço Jurídico**. Joaçaba, v. 18, n. 3, p. 637-660. 2017, p. 641. Disponível em: <https://repositorio.pucrs.br/dspace/handle/10923/11318>. Acesso em: 8/11/2021.

⁴⁶ DWORKIN, Ronald. **O Direito da Liberdade**: a leitura mora da constituição norte-americana. Tradução Marcelo Brandão Cipolla. 2ªed. São Paulo: Editora WMF Martins Fontes, 2019, p. 318-319.

⁴⁷ Nesse sentido SIMÃO e RODOVALHO, sintetizam: “Nesse ponto, é possível arrolar duas justificativas gerais para a catalogação da liberdade de expressão como direito fundamental: uma de natureza instrumental e outra constitutiva (moral). O objetivo instrumental para a proteção da liberdade de expressão faz alusão aos benefícios de que goza a sociedade com a permissão que as pessoas têm de poder falar o que bem entenderem. [...] A justificativa constitutiva, por sua vez, agrega um valor intrínseco à liberdade de expressão. Ela é importante não apenas por causa das consequências que provoca, mas fundamentalmente porque obriga o Estado e grupos privados a tratarem os cidadãos adultos e capazes como agentes morais responsáveis, sendo esse um traço essencial de uma sociedade política justa”. (SIMAO, José Luiz de Almeida; RODOVALHO, Thiago. A fundamentalidade do direito à liberdade de expressão: as justificativas instrumental e constitutiva para a inclusão no catálogo dos direitos e garantias fundamentais na constituição federal de 1988. **Cadernos do Programa de Pós-Graduação em Direito PPGD/UFRGS**. Edição Digital. v. 12, n. 1. 2017, p. 203-229, p. 213).

liberdade de expressão é um direito fundamental em razão de possibilitar a participação do cidadão na democracia; e, por fim, d) a liberdade de expressão é um direito fundamental em razão de uma suspeita do governo, já que os governos possuem interesses em temer pelos impactos das ideias, havendo, naturalmente, intenções fortes em repreendê-las.⁴⁸

Neste estudo, Barendt realiza um comparativo da liberdade de expressão na Inglaterra, EUA, Alemanha e sob a Convenção Europeia de Direitos Humanos (CEDH), além de analisar quatro teorias sobre a liberdade de expressão: a primeira delas está relacionada à importância de descobrir a verdade, sendo, pois “argumento da verdade” reflete a ideia de Jonh Stuart Mill de que a verdade é um bem fundamental apurando-se ao longo do tempo, mas não sem uma discussão aberta de fatos e julgamentos, independentemente de sua veracidade.

Enquanto, a liberdade de expressão, como aspecto da autorrealização, visualiza um aspecto integral do direito de cada indivíduo ao autodesenvolvimento e à realização. Essa teoria acentua o valor aos indivíduos servidos pela liberdade de expressão (autodeterminação, autonomia), em vez do bem à sociedade como um todo, protegido pelos argumentos da verdade ou da democracia.

O argumento da participação cidadã é o raciocínio mais popular usado nas democracias ocidentais modernas, cuja definição é o objetivo final da liberdade de expressão como preservação da democracia, sustentando o acesso individual ao debate público desinibido sobre questões políticas: sem discussão, a opinião pública não pode se formar de maneira significativa.⁴⁹

Os argumentos sobre a proteção da liberdade de expressão possuem natureza positiva; a quarta teoria “*suspicion of government*”, por sua vez, preceitua a existência de razões particularmente fortes para suspeitar do governo neste contexto, entendendo-se como um argumento negativo à medida em que destaca os males da regulação, em vez do bem da liberdade de expressão.

Acrescente-se, ainda, nessa linha de sistematização, as quatro categorias apontadas por Thomas I. Emerson quanto aos valores necessários para manutenção de um sistema de proteção da liberdade de expressão, são: “a) assegurar a satisfação individual; b) um meio para alcançar a verdade; c) um método de assegurar a participação da sociedade nas decisões sociais e políticas; e, d) como manutenção da balança entre a estabilidade e a mudança da sociedade.”⁵⁰

⁴⁸ BARENDT, Eric. **Freedom of speech**. 2 ed. New York: Oxford University Press, 2005. p. 09-24

⁴⁹ BARENDT, Eric. **Freedom of speech**. 2 ed. New York: Oxford University Press, 2005. p. 18/21.

⁵⁰ EMERSON, Thomas I. Toward a general theory of the first amendment. **The Yale Law Journal**. v. 72, n. 877, 1963. p. 878/879.

Interessante o levantamento realizado por Richard Moon, ao destacar a necessidade de se reconhecer o caráter social da liberdade de expressão.⁵¹ É certo que os argumentos para proteger a liberdade de expressão, pareçam concentrar somente em um ou uma combinação de três valores, como suscitados acima: na verdade, democracia e autonomia individual. Moon esclarece o caráter social na atuação humana sendo raramente mencionado nos argumentos do valor da liberdade de expressão, contudo “é a premissa não declarada de cada um”, e que “cada relato é incompleto sem algum reconhecimento de que a ação individual é realizada na interação social.”⁵²

Traçadas as linhas gerais sobre as teorias/argumentos entendidas pela doutrina como essenciais para justificar a liberdade de expressão, em sentido amplo, como um direito fundamental, cumpre, agora, analisá-las individualmente, organizando cada tópico relacionando à liberdade de expressão, respectivamente, como um direito fundamental: em razão de permitir o avanço do conhecimento e possibilitar a descoberta da verdade (*advancing knowledge and discovering truth*); em razão de assegurar uma autossatisfação individual (*assuring individual selffulfillment*); em razão de representar a forma de garantir a democracia (*provid for participation in decisionmaking by all members of society*); e, em razão de determinar a manutenção da balança entre a estabilidade e a mudança da sociedade.

1.2.1. A busca da verdade e o Mercado livre de Ideias

As primeiras lições desta teoria são atribuídas a Jonh Milton, ao apresentar sua defesa da liberdade de imprensa ao parlamento inglês, em 1644, após ter uma obra de sua autoria proibida de circular. Na famosa obra intitulada *Aeropagítica*, Milton sustenta que “a censura é um produto da Inquisição católica, e como tal, contrária ao pensamento da Inglaterra

⁵¹ Nesse sentido, Richard MOON aponta que: “Tornamo-nos indivíduos capazes de pensar e julgar, florescemos como pessoas racionais e sentimentais, quando nos juntamos à conversa com os outros e participamos da vida da comunidade. A emergência social da agência humana e da identidade individual pode ser expressa na linguagem da verdade/conhecimento, autorrealização/autonomia individual ou autogoverno demográfico. Cada relato da liberdade de expressão representa uma perspectiva particular ou dimensão da constituição da agência humana na vida da comunidade.” Tradução livre do original: “We become individuals capable of thought and judgement, we flourish as rational and feeling persons, when we join in conversation with others and participate in the life of the community. The social emergence of human agency and individual identity can be expressed in the language of truth/knowledge, individual self-realisation/autonomy, or demographic self-government. Each account of freedom of expression. represents a particular perspective on, or dimension of, the constitution of human agency in community life”. (MOON, Richard. The social character of freedom of expression. **Amsterdam Law Forum**. 2:1. 2009, p. 43-48, p. 43).

⁵² MOON, Richard. The social character of freedom of expression. **Amsterdam Law Forum**. 2:1. 2009, p. 43-48, p. 43.

protestante”.⁵³ Para o autor, tal sistemática somente poderia se tornar eficaz, caso fosse aplicada em outras áreas de conhecimento e comunicação, sob a pecha de permitir a proliferação de discursos nocivos por outros meios que não os livros.⁵⁴

Demonstra, ainda, utilizando argumentos ligados à religião, que a “virtude não significa o fato de não conhecer o pecado ou o que é proibido, mas sim o fato de conhecê-lo, e mesmo assim, escolher conscientemente por seguir outro caminho.”⁵⁵ Nesse raciocínio, prenuncia ser “impossível tornar as pessoas virtuosas pela coerção externa, já que o combate à corrupção moral se faz com o poder da escolha racional.”⁵⁶

Um dos pensamentos centrais da obra *Aeropagítica*, é a noção de interligação entre o bem e o mal e a escolha de um não eliminaria o outro. Com isso, a cognição da verdade exsurgiria do diálogo entre “o que existe de bom e de mau livros, cabendo ao leitor buscar o que neles mais lhe agradar.”⁵⁷ Nesta obra, Milton inicia a argumentação em torno da verdade, partindo das mesmas premissas estabelecidas sobre a virtude.

Tais conhecimentos contribuíram para o desenvolvimento da teoria de outro expoente, Jonh Stuart Mill⁵⁸, autor da célebre obra *On Liberty*,⁵⁹ apresentando sua ideia de proteger a liberdade de pensamento e debate, para tanto, entendendo que, caso haja, a repressão do exercício desta liberdade revelará um grande prejuízo, com três possíveis desdobramentos, o primeiro de ser possível a opinião emitida estar totalmente verdadeira, o segundo, estar parcialmente verdadeira e, finalmente, o terceiro, de estar totalmente errada. Extraí- -se do entendimento de Mill, como decorrência lógica, se a opinião for correta e o Estado entender por censurar, nesse caso, sem dúvida quem ficará com o maior prejuízo não será o autor da

⁵³ MILTON, Jonh. **Aeropagítica. Discurso pela liberdade de imprensa ao Parlamento da Inglaterra**. Rio de Janeiro: TopBooks, 1999. p. 9.

⁵⁴ Nesse sentido Milton, verbera que: “Vendo, portanto, que livros como esse e outros, abundantes, que são os mais suscetíveis de contaminar tanto a vida quanto a doutrina, não poderiam ser eliminados sem prejuízo para o saber e para o debate das ideias; e que os livros dessas duas espécies contaminam primeiro os letrados, de quem se transmite prontamente ao povo o que neles se contém de herético ou dissoluto; e que o mal se aprende perfeitamente sem livros, de mil outras maneiras que não podem ser contidas; e que a falsa doutrina se propaga mesmo sem livro, bastando-lhe apenas um professor, o qual pode não escrever e, assim, invalidar o obstáculo a qualquer proibição. Não sou capaz de revelar como o cauteloso trabalho de censurar pode eximir-se do rol das tentativas impraticáveis e vãs”. (MILTON, Jonh. **Aeropagítica. Discurso pela liberdade de imprensa ao Parlamento da Inglaterra**. Rio de Janeiro: TopBooks, 1999. p. 99).

⁵⁵ MILTON, Jonh. **Aeropagítica. Discurso pela liberdade de imprensa ao Parlamento da Inglaterra**. Rio de Janeiro: TopBooks, 1999. p. 19.

⁵⁶ MILTON, Jonh. **Aeropagítica. Discurso pela liberdade de imprensa ao Parlamento da Inglaterra**. Rio de Janeiro: TopBooks, 1999. p. 19.

⁵⁷ MILTON, Jonh. **Aeropagítica. Discurso pela liberdade de imprensa ao Parlamento da Inglaterra**. Rio de Janeiro: TopBooks, 1999. p. 17.

⁵⁸ Jonh Stuart Mill (1806-1873) foi o mais importante filósofo inglês do Século XIX. Na sua obra “On Liberty” (Da Liberdade), de 1859, Mill oferece quatro argumentos a favor da liberdade de pensamento e expressão, entendimento que resultou e, ainda, produz fortes impactos a respeito do tema em análise.

⁵⁹ Em língua portuguesa: MILL, John Stuart. **Sobre a liberdade**. Tradução de Ari R. Tank Brito. São Paulo: Hedras, 2010.

manifestação censurada, mas aqueles cujo entendimento seja antagônico, por serem restringidos de, no debate, alterar o erro pela verdade.⁶⁰

Baseados nos argumentos de Mill, a categoria do mercado de ideias surge exatamente onde o liberalismo econômico possui força e influência, ou seja, nos Estados Unidos. Essa teoria abrange a proteção constitucional para todas as expressões as quais auxiliem na compreensão adequada do mundo e da realidade, sem se preocupar se são comunicações de natureza política ou não. Deste modo, a teoria produziu resultados práticos e dogmáticos significativos para a liberdade de expressão na jurisdição deste país.

A noção de mercado livre de ideias possui estreita ligação com a finalidade da busca da verdade, conforme se denota dos escritos de Oliver Wendell Holmes, a quem é atribuída a autoria de sua formulação. No voto dissidente exarado no caso *Abrams v. United States*,⁶¹ o magistrado cunhou com primazia seu argumento quanto à perseguição existente contra a liberdade de expressão e opiniões: “Se você não tem dúvidas de suas premissas ou de seu poder e deseja um certo resultado de todo o coração, naturalmente expressa seus desejos na lei e elimina toda oposição”, e afirmou que os homens:

[...] podem vir a acreditar ainda mais do que acreditam nos próprios fundamentos de sua própria conduta que o bem final desejado é mais bem alcançado pelo livre comércio de ideias – que o melhor teste da verdade é o poder do pensamento de se fazer aceitar na competição do mercado, e essa verdade é o único fundamento sobre o qual seus desejos podem ser realizados com segurança. Essa, de qualquer forma, é a teoria da nossa Constituição. É um experimento, como toda vida é um experimento. [...] Embora esse experimento seja parte de nosso sistema, acho que devemos estar eternamente vigilantes contra tentativas de impedir a expressão de opiniões que detestamos e acreditamos serem carregadas de morte, a menos que ameacem tão iminentemente uma interferência imediata nos propósitos legais e urgentes da lei que uma verificação imediata é necessária para salvar o país.⁶²

⁶⁰ Nesse sentido, importante apontar a advertência de Jónatas Machado, afirmando “que alguns autores questionam/criticam a ideia apresentada por Milton e desenvolvida por Mill, de que a verdade sempre prevalecerá sobre o erro, bem como aduz que outros questionam a própria noção de verdade e a possibilidade de se apreender qualquer tipo de verdade pela razão”. (MACHADO, Jónatas Eduardo Mendes. **Liberdade de Expressão: dimensões constitucionais da esfera pública no sistema social**. Coimbra: Coimbra Editora, 2002. p. 241 e seguintes.)

⁶¹ ESTADOS UNIDOS DA AMÉRICA, Suprema Corte. **Abrams v. United States**, 1919. Disponível em: http://www.law.cornell.edu/supct/html/historics/USSC_CR_0250_0616_ZD.html. Acesso em 15 de dezembro de 2021.

⁶² Tradução livre do original: “[...] may come to believe even more than they believe the very foundations of their own conduct that the ultimate good desired is better reached by free trade in ideas—that the best test of truth is the power of the thought to get itself accepted in the competition of the market, and that truth is the only ground upon which their wishes safely can be carried out. That at any rate is the theory of our Constitution. [...] While that experiment is part of our system I think that we should be eternally vigilant against attempts to check the expression of opinions that we loathe and believe to be fraught with death, unless they so imminently threaten immediate interference with the lawful and pressing purposes of the law that an immediate check is required to save the country”. (ESTADOS UNIDOS DA AMÉRICA, Suprema Corte. **Abrams v. United States**, 1919. Disponível em: http://www.law.cornell.edu/supct/html/historics/USSC_CR_0250_0616_ZD.html. Acesso em 15 de dezembro de 2021.)

O aludido julgamento, deu-se num contexto no qual a liberdade de expressão era comumente censurada e manipulada pelo governo, sobretudo em defesa dele. Tanto que fora editada a chamada Lei de espionagem (*Espionage Act*, 1917), servindo de fundamento da condenação dos russos, no caso supracitado, acusados de “conspirar para violar disposições da Lei de Espionagem do Congresso (seção 3, título I, da Lei 15 de junho de 1917, c. 30 , 40 *Stat.* 219 , conforme alterado pela Lei de 16 de maio de 1918, c. 75, 40 *Stat.* 553 [Comp. St. 1918, § 10212c]).”⁶³ Essa Lei foi objeto de alterações, “sendo que uma das principais ocorreu em 1918, com a aprovação do Ato de Sedição (*Seditious Act*), inaugurando uma imunidade do governo americano às críticas.”⁶⁴

Nesse sentido, o caso *Abrams v. United States* serviu para delinear alguns paradigmas da liberdade de expressão, uma vez que, com a criminalização das críticas, tornou-se proibida a própria comunicação em si. Jacob Abrams⁶⁵ e demais réus, sustentavam, em síntese, que seus atos não eram ilegais, pois estariam sob a égide da liberdade de expressão e de imprensa, conforme previsto na Primeira Emenda da Constituição dos Estados Unidos, bem como a Lei de Espionagem era integralmente inconstitucional, porque confrontava os ditames da referida emenda, contudo, restaram condenados.

Posteriormente, a Suprema Corte firmou entendimento sobre a inconstitucionalidade da Lei de Espionagem sustentando: “suficientemente discutida e definitivamente negada em *Schenck v. Estados Unidos* e *Baer v. Estados Unidos*, 249 US 47, 39 Sup. Ct. 247, 63 L. Ed. 470, e em *Frohwerk v. Estados Unidos*, 249 US 204 , 39 Sup. Ct. 249, 63 L. Ed. 561.”⁶⁶

Nesse julgado, o Juiz Holmes propôs uma teoria para justificar a defesa da liberdade de expressão: a teoria da verdade, baseada em um experimento social dentro de um livre

⁶³ ESTADOS UNIDOS DA AMÉRICA, Suprema Corte. **Abrams v. United States**, 1919. Disponível em: http://www.law.cornell.edu/supct/html/historics/USSC_CR_0250_0616_ZD.html. Acesso em 15 de dezembro de 2021.

⁶⁴ LAURENTIIS, Lucas Catib de; THOMAZINI, Fernanda Alonso. Liberdade de expressão: Teorias, Fundamentos e Análise de Casos. **Revista Direito e Práxis**. Rio de Janeiro, V. 11, N. 4, 2020, p. 2260-2301, p. 2262/2263.

⁶⁵ “Jacob Abrams, juntamente com outros quatro imigrantes russos, foram presos e condenados por violarem o Ato de Sedição de 1918. Seu crime foi distribuir pelas ruas de Nova Iorque panfletos que criticavam os EUA por enviarem tropas para a Rússia, além de incentivarem os trabalhadores da indústria de munições a fazer uma greve em protesto contra a campanha norte-americana na primeira guerra. Após a sua condenação em primeira instância, os réus apelaram, baseando sua defesa na primeira emenda, sob o argumento de que seu direito fundamental à liberdade de expressão havia sido desrespeitado. Alegaram que os atos por eles cometidos seriam irrelevantes, pois sua conduta estaria protegida pelo direito à liberdade de expressão. A própria Lei de sedição, diziam, deveria ser considerada completamente inconstitucional.” (LAURENTIIS, Lucas Catib de; THOMAZINI, Fernanda Alonso. Liberdade de expressão: Teorias, Fundamentos e Análise de Casos. **Revista Direito e Práxis**. Rio de Janeiro, v. 11, n. 4, 2020, p. 2260-2301, p. 2263).

⁶⁶ Tradução livre do original: “This contention is sufficiently discussed and is definitely negated in *Schenck v. United States* and *Baer v. United States*, 249 U. S. 47, 39 Sup. Ct. 247, 63 L. Ed. 470, and in *Frohwerk v. United States*, 249 U. S. 204, 39 Sup. Ct. 249, 63 L. Ed. 561.” Disponível em: <https://www.law.cornell.edu/supremecourt/text/250/616>. Acessado última vez em 17/02/2022.

mercado de ideias. Nessa transação de ideias e modos de compreender o mundo, o valor da verdade de um enunciado “deve ser aferido pelo resultado do confronto da ideia nele contida com as ideias que lhe são contrárias.”⁶⁷ Assim, o valor da verdade é definido pela livre troca de ideias. Destarte para uma ideia atingir impacto e alcançar valor nesse mercado, “ela deve ser o resultado de reflexão coletiva, pois uma verdade autodeclarada nada mais será do que uma falsa verdade.”⁶⁸

Alexander sintetiza a teoria do livre mercado de ideias surgindo a partir da percepção de “que a liberdade para disseminar novas informações e para criticar visões prevalecentes é necessária para eliminar concepções equivocadas sobre fatos e valores.”⁶⁹

Por seu turno, Chequer afirma que a liberdade de expressão, segundo essa perspectiva, não é compreendida apenas como um direito individual, mas também como um bem social. Nesse sentido, a teoria firma uma premissa no sentido de o “mais íntegro e racional julgamento só pode ser alcançado se considerarmos todos os fatos e argumento que podem ser colocados em favor ou contra uma proposição.”⁷⁰

Deste modo, a aplicação da teoria abrange um sistema de considerar todas as opiniões, para a formação de outra(s), sem a conformação de uma verdade estática, partindo do pressuposto que vários aspectos devem serem conhecidos por aquele que busca se apropriar da verdade. Devendo considerar todas as alternativas, provando seu julgamento à medida de estar exposto às teses contrárias, elaborando todos os usos de distintos pensamentos, para ter condições de separar o verdadeiro do falso; a verdade, da mentira.

Franklyn Haiman entende que o mercado de ideias não deve ser desfeito, devendo haver uma ampla defesa deste, pois, para o autor, não é crível que, “numa sociedade livre, discursos que possam ser considerados nocivos devam ser punidos pela lei.”⁷¹ Defende, ainda, que a solução sempre deve ser “mais expressão” (*more speech*), sendo assim, deve ser

⁶⁷ LAURENTIIS, Lucas Catib de; THOMAZINI, Fernanda Alonso. Liberdade de expressão: Teorias, Fundamentos e Análise de Casos. **Revista Direito e Práxis**. Rio de Janeiro, v. 11, n. 4, 2020, p. 2260-2301, p. 2264/2265

⁶⁸ LAURENTIIS, Lucas Catib de; THOMAZINI, Fernanda Alonso. Liberdade de expressão: Teorias, Fundamentos e Análise de Casos. **Revista Direito e Práxis**. Rio de Janeiro, v. 11, n. 4, 2020, p. 2260-2301, p. 2264/2265.

⁶⁹ ALEXANDER, Larry. **Is There a Right of Freedom of Expression?** New York: Cambridge University Press, 2005. p.128.

⁷⁰ CHEQUER, Cláudio Márcio de Carvalho. **A liberdade de expressão como direito fundamental preferencial *prima facie*** (análise crítica e proposta de revisão ao padrão jurisprudencial brasileiro). 2010. 340 f. Tese de Doutorado - Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2010. p. 31.

⁷¹ HAIMAN, Franklyn S. **“Speech Acts” and the First Amendment**. Carbondale: Southern Illions University Press, 1993. p. 34

estimulado ainda mais o debate, principalmente com aqueles que se sentiram ofendidos, para que se expressem de forma oposta ao discurso ofensor.⁷²

Nessa perspectiva teórica, quando uma discussão é mantida em aberto, não se deve concluir, necessariamente, como verdadeira, pois, no primeiro contato, transparecesse ser. Esse processo se aplica, indistintamente, a qualquer opinião, independentemente de ser falsa ou pernicioso. Uma opinião inaceitável pode ser verdadeira ou parcialmente verdadeira. Ademais, ainda que a nova opinião seja inteiramente falsa, a apresentação e abertura de debate a respeito representa um vital propósito social.⁷³

A teoria da busca da verdade e o mercado livre de ideias possuem relações íntimas, por partirem da ideia de verdade. Para Jónatas Machado, persiste a diferença entre ambas, por dois motivos: primeiro porque a verdade tem sofrido historicamente interferências, como, por exemplo, o crescimento do conceito de verdade, apresentando uma certa dose de flexibilização em relação à fidedignidade, ou seja, a verdade em si. Percebeu-se então “a sobrevivência e a coexistência de diferentes verdades, contraditórias entre si, traduzindo diferentes consensos”⁷⁴, de maneira que “uma ideia legitima-se, não através da sua correspondência substantiva com critérios objetivos pré-determinados, mas por ser o resultado, ainda que provisório de um procedimento dialógico e crítico livre e aberto.”⁷⁵

O segundo motivo apontado por Machado são as teorias em relevo recebendo várias críticas, resumidamente, tem-se que teoria da busca pela verdade é contestada pela sua abordagem epistemológica da verdade. Já com a teoria do mercado de ideias a crítica diz respeito sobre os problemas na operação e nas imperfeições do próprio mercado.⁷⁶ Ambas as teorias se baseiam numa busca racional da verdade ou o mais próximo possível dela, por meio do confronto de ideias.

Como delineado, as teorias sobre a verdade não passam ilesas de críticas, sendo questionada por diversos autores. Uma das principais críticas está em argumentos que apontam para a desconsideração das desigualdades sociais e econômicas existentes. Em um

⁷² HAIMAN, Franklyn S. **“Speech Acts” and the First Amendment**. Carbondale: Southern Illinois University Press, 1993. p. 34.

⁷³ CHEQUER, Cláudio Márcio de Carvalho. **A liberdade de expressão como direito fundamental preferencial prima facie** (análise crítica e proposta de revisão ao padrão jurisprudencial brasileiro). 2010. 340 f. Tese de Doutorado - Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2010. p. 32.

⁷⁴ MACHADO, Jónatas Eduardo Mendes. **Liberdade de Expressão: dimensões constitucionais da esfera pública no sistema social**. Coimbra: Coimbra Editora, 2002. p. 248.

⁷⁵ MACHADO, Jónatas Eduardo Mendes. **Liberdade de Expressão: dimensões constitucionais da esfera pública no sistema social**. Coimbra: Coimbra Editora, 2002. p. 248.

⁷⁶ MACHADO, Jónatas Eduardo Mendes. **Liberdade de Expressão: dimensões constitucionais da esfera pública no sistema social**. Coimbra: Coimbra Editora, 2002. p. 248.

sistema de liberdade discursiva ampla, as verdades são definidas pelas vozes daqueles detentores do poder econômico.⁷⁷

Além disso, outra fragilidade consiste exigência da busca da verdade. A teoria exige uma deliberação racional das controvérsias e uma cooperação entre as partes interessadas, as quais precisam escutar e serem ouvidas simultaneamente, contudo, dificilmente essa interação ocorrerá sem interferências, considerando a inconstância dos seres humanos em funcionar racionalmente.

Quanto à função de buscar a verdade proposta pela teoria do livre mercado de ideias, Robert Post considera ser insuficiente sem um conjunto de práticas sociais. Adverte que, em um discurso aberto é necessário a capacidade para escutar, além do compromisso de autoavaliação, a qual, por sua vez, implica respeito mútuo e civilidade. Analisou, ainda, sobre necessidade do entendimento de verdade política construída pela sociedade num debate democrático, por meio da manifestação de suas percepções, modo de agir e sentir a realidade, não declinando tal defesa ao Governo, ou ainda, ao Judiciário. Nesse sentido, é premente a defesa do discurso das minorias, pois, tais posicionamentos são partes integrantes do livre mercado, fundamentais para o bom funcionamento desse sistema.⁷⁸

Nesse ponto, percebe-se a teoria da verdade permitindo a possibilidade do Estado regular várias formas de manifestar e expressar. Havendo, com isso, uma classificação indesejável de quais seriam as opiniões merecedoras de proteção, restando obscura quem realizaria essa classificação, e qual método seria adotado. Nesse cenário, é oportuno o questionamento de Laurentiis e Thomazini: quais são os requisitos sociais da busca pela verdade? Surgindo um distanciamento entre os resultados práticos da teoria do livre mercado de ideias e sua formação teórica, dificultando o acesso de todos a esse mercado, de forma igualitária. Por outro lado, “aquele que já tem acesso a ele tem a tendência de silenciar o seu oponente. É errado pensar que o livre mercado de ideias proteja todo o discurso que transmita alguma ideia. Como todo mercado, esse também é seletivo”.⁷⁹

Cass Sunstein entende ser discutível a teoria do livre mercado de ideias, pois sua aplicação é deveras difícil de ser implementada, por não haver condições favoráveis na realidade social, pois, na democracia existente não há disposição igualitária e firme de

⁷⁷ MOON, Richard, **The constitutional protection of freedom of expression**. Toronto: University of Toronto Press, 2000. p. 13.

⁷⁸ POST, Robert. Reconciling Theory and Doctrine in First Amendment Jurisprudence. **California Law Review**. California, v. 88, 2000, p. 2353-2374, p. 2365.

⁷⁹ LAURENTIIS, Lucas Catib de; THOMAZINI, Fernanda Alonso. Liberdade de expressão: Teorias, Fundamentos e Análise de Casos. **Revista Direito e Práxis**. Rio de Janeiro, v. 11, n. 4, 2020, p. 2260-2301, p. 2266.

informações capazes de subsidiar a formação de opiniões e decisões da sociedade, existindo a possibilidade de sua aplicação se tornar difícil, pela presença de controle e manipulação de informações, muitas vezes sem conhecimento das fontes. Sunstein visualiza um ponto positivo na teoria, o qual permite a construção de ideias por meio de um debate, proporcionando um processo deliberativo, essencial para a manutenção de qualquer democracia.⁸⁰

Portanto, apesar da teoria da verdade apresentar limitações, não deixa de ser um dos pontos essenciais para justificar a liberdade de expressão: “ideias são reguladas e combatidas por meio de outras ideias, não com poder ou força. Essa é a origem da liberdade de expressão”.⁸¹

1.2.2. Teoria da autossatisfação (*self-fulfilment*) individual

Diferentemente da teoria da verdade, a qual coloca a liberdade de expressão como um direito instrumental para maximizar o acesso à verdade, a teoria da autossatisfação ou da autonomia individual pressupõe a liberdade de expressão, tendo em vista que compõe a própria condição humana, como um direito humano, sendo necessária e valiosa por si só, possuindo assim, um valor intrínseco, um bem independente, um fim; por outro lado, o exercício da liberdade de expressão pode ser considerado como condutor para o desenvolvimento de pessoas mais maduras e reflexivas.

Nesse viés, a expressão é parte integrante do desenvolvimento de ideias, da exploração mental e da autoafirmação, representando um meio para realizar sua potencialidade no ser humano. Assim, a restrição de opinião, expressão e crença, representa uma afronta à dignidade da pessoa humana, negando a natureza essencial do homem, valor que induz a centralidade do homem em todo o ordenamento jurídico.⁸²

Farber define sinteticamente a teoria da autorrealização (ou autossatisfação) como um instrumento para a autorrealização das pessoas. A autorrealização cinge em torno de ter facilmente acesso as mais diversificadas ideias, possibilitando alternativas para construção da percepção de vida ao indivíduo. Numa visão mais ampla da liberdade de expressão, Faber

⁸⁰ SUSTEIN, Cass. **Democracy and the problem of free speech**. New York: The Free Press, 1993. p. 21 e ss.

⁸¹ LAURENTIIS, Lucas Catib de; THOMAZINI, Fernanda Alonso. Liberdade de expressão: Teorias, Fundamentos e Análise de Casos. **Revista Direito e Práxis**. Rio de Janeiro, v. 11, n. 4, 2020, p. 2260-2301, p. 2266.

⁸² CHEQUER, Cláudio Márcio de Carvalho. **A liberdade de expressão como direito fundamental preferencial *prima facie*** (análise crítica e proposta de revisão ao padrão jurisprudencial brasileiro). 2010. 340 f. Tese de Doutorado - Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2010. p. 28.

aduz que esta abarca a liberdade de escritores e artistas desempenharem suas habilidades, como forma de desenvolver uma cultura nacional, e caso haja uma limitação da expressão estar-se-ia, pois, colaborando com o enfraquecimento desta cultura.⁸³

Ronald Dworkin⁸⁴, sendo um dos defensores dessa teoria, defende ser imprescindível a existência de obediência, tanto à igualdade, como à democracia, como meio de possibilitar a participação da maior quantidade possível de pessoas na consolidação de um entendimento moral da sociedade. Desta forma, a liberdade de expressão favorece a construção de igualdade civil numa sociedade democrática.⁸⁵

O fundamento da autossuficiência, para Dworkin, possui duas dimensões. A primeira reside no fato de pessoas moralmente responsáveis assumirem a postura de se posicionarem sobre o que entendem ser o bom ou o mal, tanto na vida, como na política; além do que é verdadeiro ou falso, quanto à justiça, ou, até mesmo, à fé. A segunda dimensão está em expressar suas próprias convicções aos outros, sob uma perspectiva de respeito recíproco, e pela intenção em distribuir a verdade, além da satisfação da justiça, prezando pela vitória do bem comum.⁸⁶

Desta forma, o estímulo para as pessoas pensarem por si mesmas é um dos fatores colaboradores para a existência da autonomia e independência moral,⁸⁷ sem que haja interferência do governo. Assim, o governo deve tratar todos como pessoas independentes.

Dworkin destaca a possibilidade de os cidadãos serem agentes morais, advinda do dever de o governo tratar a todos como agentes independentes.

Corroborando com Dworkin, Levin observa que:

depende do indivíduo como, se é que, e até que ponto, ele adota as ideias apresentadas em uma sociedade que protege a liberdade de expressão, e assim, de

⁸³ FARBER, Daniel A. **The First Amendment**. 2ª ed. New York: Foundation, 2003. p. 03.

⁸⁴ Para melhor compreensão sobre o panorama do pensamento de Ronald Dworkin sobre a liberdade de expressão, conferir: ROSA, Leonardo Gomes Penteado. **O liberalismo igualitário de Ronald Dworkin: o caso da liberdade de expressão**. Dissertação. Universidade de São Paulo, Faculdade de Direito, 2014, 254f.

⁸⁵ Nesse sentido, de defesa da liberdade de expressão como meio de garantir o desenvolvimento de cada pessoa, além de contribuir para a construção de uma sociedade mais democrática, SANKIEVICZ expõe, “se o processo democrático deve garantir a cada pessoa o seu desenvolvimento e o mesmo potencial de influenciar o resultado das decisões coletivas, respeitando, ao menos no plano abstrato, autonomia e igualdade, qualquer decisão política que seja incompatível com a possibilidade de que alguém fale ou atinja um público determinado constituiria uma restrição indevida a esses valores”. (SANKIEVICZ, Alexandre. **Liberdade de expressão e pluralismo: perspectivas de regulação**. São Paulo: Editora Saraiva, 2011, p. 23/24)

⁸⁶ DWORKIN, Ronald. **O Direito da Liberdade: a leitura mora da constituição norte-americana**. Tradução Marcelo Brandão Cipolla. 2ªed – São Paulo: Editora WMF Martins Fontes, 2019, p. 319/320.

⁸⁷ Nesse contexto, a liberdade de expressão para Dworkin, se destina à proteção daquele que fala, uma vez que os indivíduos têm o direito de falar não para que os outros se beneficiem, mas porque eles mesmos sofreriam um dano ou insulto inaceitável se fossem censurados. (DWORKIN, Ronaldo. **Uma questão de princípio**. Tradução Luis Carlos Borges. 2 ed. São Paulo: Livraria Martins Fontes Editora, 2005, p. 575/576.)

acordo com o esquema de Dworkin, ao indivíduo cabe definir o quanto ele deseja ser um agente moral⁸⁸.

Desta forma, entende-se que a independência moral somente será possível, quando for opcional, sem interferência do governo na liberdade de expressão. Com isso, a sociedade continua sendo democrática, ainda que alguns indivíduos decidam por não usufruir desta liberdade.

Por outro lado, Levin possui entendimento diverso de Dworkin, quanto ao resultado desta teoria. Dworkin não aceita a realização de flexibilização da liberdade de expressão, pois, para ele, qualquer relativização estaria violando este direito fundamental. Conforme o entendimento de Levin, prestigiar a liberdade de expressão possibilita a opiniões de menor valor, como racistas, homofóbicas, sexistas, ou seja, ao exercer essa liberdade, surgem chances para a elaboração de discursos que afrontem a própria igualdade.⁸⁹

C. Edwin Baker, outro expoente dessa teoria centrada no indivíduo, adota como valores, dentre os apontados por Emerson,⁹⁰ o *self-fulfillment* (item a) e o *participation in change* (item c) e, justifica sua posição, desenvolvendo um argumento sobre a legitimidade estatal e sobre a permissão dada a maioria em restringir direitos. Assim, no regime democrático, a obediência às normas somente seria exigida caso as pessoas fossem tratadas igualitariamente e de forma independente.⁹¹

O autor sustenta o dever de obediência do sistema democrático à igualdade e à autonomia, tendo em vista que instigaria a implantação de um regime democrático, com a consequente imposição de balizas à vontade majoritária, bem como limitaria o Estado em produzir decisões opostas aos valores estipulados pela comunidade. Desta forma, a liberdade de expressão exerceria o papel de sustentáculo dos valores cruciais para a sociedade democrática, dentre eles a importância do desenvolvimento humano e as conquistas de cada indivíduo, e da proteção da legitimidade do sistema democrático.

Em sua obra, Baker indica exemplos sobre a relevância da liberdade de expressão no desenvolvimento do indivíduo, afirmando que protestar contra uma guerra, não configura um

⁸⁸ LEVIN, Abigail. Pornography, Hate Speech and their challenge to Dworkin, **Public Affairs Quarterly**, v. 23, n. 4, p. 357-373, 2009, p. 359. Disponível em: <https://www.jstor.org/stable/40441541>. Acesso em: 10/02/2022.

⁸⁹ LAURENTIIS, Lucas Catib de; THOMAZINI, Fernanda Alonso. Liberdade de expressão: Teorias, Fundamentos e Análise de Casos. **Revista Direito e Práxis**. Rio de Janeiro, v. 11, n. 4, 2020, p. 2260-2301, p. 2269/2270.

⁹⁰ “a) assegurar a satisfação individual; b) como um meio para alcançar a verdade; c) como um método de assegurar a participação da sociedade nas decisões sociais e políticas; e, d) como manutenção da balança entre a estabilidade e a mudança da sociedade”. (EMERSON, Thomas I. Toward a general theory of the first amendment. **The Yale Law Journal**. 72:877 (1963). p. 878/879)

⁹¹ BAKER, C. Edwin. **Human Liberty and Freedom of Speech**. New York: Oxford University Press, 1989. p. 278.

discurso e não é totalmente político, apesar de tratar de um problema de interesse coletivo, pois, o cidadão, ao manifestar sua opinião num protesto, não tem necessariamente poder de influência na permanência ou não do conflito armado, ou na própria tomada de decisão das autoridades governantes.

A atitude do manifestante, durante o protesto, pouco tem a ver com sua expectativa de afetar a continuidade da guerra ou mesmo com o objetivo de ser ouvido pelos seus governantes.

Desta forma, Baker indica caminhos para maximizar os efeitos da liberdade de expressão para outras condutas do indivíduo, além do discurso meramente político, aplicando também para outras formas de discurso.

Nesse ponto, o autor encontra convergência com o pensamento de Martin Redish, contudo, divergindo quanto ao valor-chave adotado, pois, para Redish, o valor a ser perseguido pela liberdade de expressão é a autorrealização individual, ou seja, a capacidade do indivíduo desenvolver suas habilidades, dentre elas a de controlar suas próprias decisões.⁹² Consta-se o fato dessa teoria, segundo Redish, possuir, como premissa, a autorrealização individual representar um valor relevante, não apenas da liberdade de expressão, mas também do sistema democrático.

Ademais, Redish sustenta existirem, não apenas opiniões políticas, mas também, opiniões apolíticas, essas podendo influenciar para alcançar um ideal de auto autorrealização (*self-realization*). Nesse raciocínio, a liberdade de expressão serviria para tomadas de decisões, as quais influenciariam na vida dos indivíduos, em setores, para além da política. Assim ambos os discursos estariam sob a égide da liberdade de expressão, independentemente de eventual classificação como discurso político ou apolítico, merecendo ambos, portanto, proteção do sistema jurídico.⁹³

Por fim, cumpre fincar que várias críticas são endereçadas a essa teoria, por não explicar: o porquê de o indivíduo autor do discurso dever ter prioridade em relação aos seus destinatários; a possibilidade de produzir um efeito silenciador, pois, o acesso diferenciado aos meios de comunicação de massa permite a grupos hegemônicos poder de monopólio do discurso e silenciando os menos favorecidos; e não justifica o fato do discurso receber

⁹² REDISH, Martin H. The Value of Free Speech. **University of Pennsylvania Law Review**, v. 130, n. 3, p. 591-645, 1982. p. 593 Disponível em: <https://www.jstor.org/stable/3311836?refreqid=excelsior%3A067a6da4e7fc482a0901ef01532bd0f1>. Acesso em 13/12/2021.

⁹³ REDISH, Martin H. The Value of Free Speech. **University of Pennsylvania Law Review**, v. 130, n. 3, p. 591-645, 1982. p. 603 e seguintes. Disponível em: <https://www.jstor.org/stable/3311836?refreqid=excelsior%3A067a6da4e7fc482a0901ef01532bd0f1>. Acesso em 13/12/2021.

proteção maior em relação a um grande número de atividades também essenciais para o desenvolvimento da personalidade.⁹⁴

Nessa linha, Chequer resume os principais argumentos das críticas:

- a) qual a razão para a proteção específica da liberdade de expressão e não como pedem os libertários, de outras condutas igualmente dedutíveis da autorrealização?
- b) em que medida se pode restringir a liberdade de expressão em diversas situações concretas?
- c) que fazer quando determinada expressão, com base no valor da autorrealização, põe em causa a dignidade e o desenvolvimento da personalidade de outrem?
- d) qual o peso deverá ser atribuído, em concreto, aos direitos dos editores, dos distribuidores e de todos os demais participantes no processo comunicativo?⁹⁵

Seguindo o argumento de Canotilho e Machado, muitas dessas perguntas se esvaziariam, com a afirmação de o direito fundamental a liberdade de expressão também se aplicar aos demais direitos, liberdades e garantias fundamentais e de a liberdade ser a regra e a restrição ser a exceção.⁹⁶

1.2.3. Teoria do autogoverno ou autodeterminação democrática

Essa teoria compreende um dos mais importantes pilares de sustentação do conceito de liberdade de expressão, possuindo a capacidade de unificá-lo ao conceito de democracia e servindo como um importante esteio constitucional do Estado de Direito Moderno, delimitando a liberdade de expressão como um instrumento de autogoverno.

Nesse sentido, a liberdade de expressão passou a ser concebida como uma consequência do sistema democrático de tomada de decisões e não como uma forma de autorrealização, na medida da contribuição para a formação de opinião pública acerca dos debates públicos.⁹⁷

Como afirma Robert Dahl, é essencial numa democracia a liberdade de expressão e a multiplicidade de fontes de informação. Nesse sentido, é indispensável assegurar a participação dos cidadãos na vida pública, como forma de influenciar na tomada de decisões, além de compreender os movimentos e decisões políticas governamentais.⁹⁸

⁹⁴ SANKIEVICZ, Alexandre. **Liberdade de expressão e pluralismo: perspectivas de regulação**. São Paulo: Editora Saraiva, 2011, p. 24/25.

⁹⁵ CHEQUER, Cláudio Márcio de Carvalho. **A liberdade de expressão como direito fundamental preferencial prima facie** (análise crítica e proposta de revisão ao padrão jurisprudencial brasileiro). 2010. 340 f. Tese de Doutorado - Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2010. p. 29/30.

⁹⁶ CANOTILHO, J. J. Gomes; MACHADO, Jónatas E. M. **“Reality Shows” e a Liberdade de Programação**. Coimbra: Coimbra Editora, 2003. p. 12.

⁹⁷ SANKIEVICZ, Alexandre. **Liberdade de expressão e pluralismo: perspectivas de regulação**. São Paulo: Editora Saraiva, 2011, p. 29-32.

⁹⁸ DAHL, Robert. **La democracia**. Uma guia para los cuidadanos. Taurus, 1999. p. 112.

Machado⁹⁹ explicita as ideias de Alexander Meiklejohn, Cass Sunstein, Robert Bork e, Zecharia Chaffee,¹⁰⁰ como os principais pensadores do entendimento sobre a autodeterminação democrática. Como visto, um dos importantes representantes dessa corrente é Alexander Meiklejohn, ele realiza um estudo aprofundado sobre o *self-government*, no qual o autogoverno abarca o conceito de governo do povo e pelo povo; em sua investigação, descarta a possibilidade de os indivíduos retirarem o poder do Estado por não atender os anseios da população. Para o autor, tendo em vista o rumo do governo estar diretamente ligado às decisões e ao bem comum dos cidadãos, não deve haver espaço para a supressão de ideias e posicionamentos diversos, a fim de evitar o empobrecimento do debate público.¹⁰¹

Nesse sentido, Meiklejohn exemplifica os parlamentares, em sua essência como representantes do povo, motivo pelo qual detém imunidade de opinião; todavia, imperativo afirmar que a finalidade da proteção não se destina apenas àqueles em exercício de atividades governamentais, devendo a sua extensão ser aplicada ao povo em sua totalidade¹⁰². Desta forma, ao analisar a posição de Meiklejohn, Laurentiis e Thomazini afirmam que a “legitimidade de um governo democrático vem do povo e do fato de o poder exercido pelo governo pertencer aos governados, sendo que, esse exercício deve ocorrer sempre com base no consentimento destes”¹⁰³, ressaltando-se a necessidade de consentimento, não de submissão.

Conforme verbera Meiklejohn: “homens livres não são não-governados. Eles são governados – por si mesmos.”¹⁰⁴ Devendo ficar claro, entretanto, que esse autogoverno não configura uma licença para que os cidadãos descumpram as leis, das quais discordam, sob o risco de iniciar caos e anarquia.

Importante destacar a crítica de Jónatas Machado sobre a teoria da autodeterminação democrática devido ao enfraquecimento da liberdade de expressão sobre opiniões que não

⁹⁹ MACHADO, Jónatas Eduardo Mendes. **Liberdade de Expressão: dimensões constitucionais da esfera pública no sistema social**. Coimbra: Coimbra Editora, 2002. p. 255.

¹⁰⁰ Conferir: CHAFFEE, Zechariah Jr. **Freedom of Speech**. 1920.

¹⁰¹ MEIKLEJOHN, Alexander. **Free Speech and its Relation to Self-Government**. New York: Harper & Brothers Publishers, 1948, p. 26 e seguintes.

¹⁰² MEIKLEJOHN, Alexander. **Free Speech and its Relation to Self-Government**. New York: Harper & Brothers Publishers, 1948, p. 36.

¹⁰³ Nesse ponto, Laurentiis e Thomazini, dissertam que, “Meiklejohn observa que a submissão vem da obediência de um poder externo. Ela decorre do medo e falta de alternativa do destinatário da ordem, que não tem opção a não ser seguir o comando. É o caso da submissão de um escravo, que apenas obedece às regras impostas pelo seu senhor. Já o consentimento não é imposto. Ele decorre de uma escolha: ceder um poder que é seu. Nesse caso, não há imposições, mas sim em direitos. O povo que consente em ser governado tem o direito e obrigação de cobrar uma boa gestão”. (LAURENTIIS, Lucas Catib de; THOMAZINI, Fernanda Alonso. *Liberdade de expressão: Teorias, Fundamentos e Análise de Casos*. **Revista Direito e Práxis**. Rio de Janeiro, v. 11, n. 4, 2020, p. 2260-2301, p. 2272)

¹⁰⁴ MEIKLEJOHN, Alexander. **Free Speech and its Relation to Self-Government**. New York: Harper & Brothers Publishers, 1948, p. 16.

são, necessariamente, fomentadoras das demandas de interesse público; sobre isso, esclarece-se, somente, a fragilização do resguardo às diversas linguagens de expressão, em quaisquer áreas, tendo em vista o enfoque protetivo ser dado, apenas, aos discursos políticos ou de interesse público.¹⁰⁵

Mister destacar a dicotomia sobre os métodos de proteção constitucional no sistema jurídico norte-americano apresentado por Meiklejohn: *freedom of expression* e o *liberty of expression*. O primeiro é protegido pela Primeira Emenda da Constituição dos Estados Unidos¹⁰⁶, o segundo, por outro lado, recebe a proteção da Quinta Emenda da Constituição dos Estados Unidos¹⁰⁷.

A *freedom of expression* (liberdade de expressão) possui como principal escopo proteger opiniões, discursos e falas de interesse coletivo, não sendo possível sofrer nenhum tipo de supressão. Por sua vez, a *liberty of expression* preceitua a proteção ao discurso privado, equiparando-se, inclusive, à propriedade; desta forma, pode ser submetida ao devido processo legal, sob a argumentação da sua imprescindibilidade e, conseqüentemente, a autorização da sua supressão.¹⁰⁸

Outro autor que aborda a teoria do autogoverno é Cass Sunstein, o qual imerge na concepção do conceito de soberania e democracia.¹⁰⁹ Assim como Meiklejohn e Sunstein, também possui um entendimento de que a liberdade de expressão possui facetas distintas, desafiando, portanto, os métodos jurídicos adequados ao padrão em análise. Sunstein denomina de *Two-Tier First Amendment*, pela tradução, duas camadas da Primeira Emenda.

A primeira camada dialoga sobre o discurso político; segundo o autor, este constitui o núcleo duro (*core*) do resguardo constitucional à liberdade de expressão. Em razão da importância, essa proteção demanda um alto valor, o qual é chamado de *high value*.¹¹⁰ A

¹⁰⁵ MACHADO, Jónatas Eduardo Mendes. **Liberdade de Expressão**: dimensões constitucionais da esfera pública no sistema social. Coimbra: Coimbra Editora, 2002. p. 262

¹⁰⁶ Diz a primeira emenda que: “O Congresso não fará nenhuma lei a respeito do estabelecimento de uma religião, ou proibindo seu livre exercício; ou cerceando a liberdade de expressão ou de imprensa; ou o direito do povo de se reunir pacificamente e de solicitar ao Governo a reparação de queixas.” Disponível em http://www.archives.gov/exhibits/charters/bill_of_rights_transcript.html. Acesso em 28/12/2021.

¹⁰⁷ Diz a quinta emenda que: “Ninguém será detido para responder por crime capital, ou outro crime infame, a não ser sob apresentação ou acusação de um Grande Júri, exceto em casos ocorridos nas forças terrestres ou navais, ou na Milícia, quando em serviço efetivo em tempo de Guerra ou perigo público; nem qualquer pessoa será sujeita pelo mesmo delito a ser duas vezes colocada em risco de vida ou integridade; nem será obrigado em qualquer processo criminal a servir de testemunha contra si mesmo, nem ser privado da vida, liberdade ou propriedade, sem o devido processo legal; nem a propriedade privada será tomada para uso público, sem justa compensação”. Disponível em http://www.archives.gov/exhibits/charters/bill_of_rights_transcript.html. Acesso: em 28/12/2021.

¹⁰⁸ MEIKLEJOHN, Alexander. **Free Speech and its Relation to Self-Government**. New York: Harper & Brothers Publishers, 1948, p. 38.

¹⁰⁹ SUNSTEIN, Cass R. **Democracy and the Problem of Free Speech**. New York: Free Press, 1995. p. xvi.

¹¹⁰ SUNSTEIN, Cass R. **Democracy and the Problem of Free Speech**. New York: Free Press, 1995. p. 122.

segunda camada, por sua vez, refere-se aos discursos não políticos (*nonpolitical speech*); portanto, eles recebem um tratamento com menos densidade, cuja localização é dada fora do núcleo duro e garantista da Constituição Americana, consequentemente, taxado de baixo valor (*low value*). Em resumo, Sunstein privilegia a proteção constitucional aos discursos políticos, possibilitando que alguns discursos (não políticos) possam ser restringidos.

Owen M. Fiss, em linha teórica parecida com a de Meiklejohn e Sunstein, disserta sobre a Primeira Emenda da Constituição dos Estados Unidos depender de uma interpretação assimétrica, a qual traz a faculdade em proteger todos os tipos de discursos, mesmo o denominado discurso de ódio (*hate speech*).¹¹¹

Nota-se, desta forma, uma certa proximidade do entendimento de Fiss ao de Meiklejohn e Sunstein, principalmente, quanto à ideia de autodeterminação coletiva. Contudo, é possível perceber diferenças a partir da menção de Fiss sobre o “efeito silenciador do discurso”, no qual trata a concepção norte-americana sobre a liberdade de expressão, sobre o aspecto libertário e outro democrático. Ademais, afirma: “A sociedade é mais que um encontro em praça pública, e o Estado significativamente mais que um mediado”,¹¹² nessa visão, tanto pela autonomia privada, quanto pela tutela estatal. Em algum aspecto, a liberdade de expressão é confrontada por grupos que estão à margem do debate público. Nessa perspectiva, Fiss aponta para a autodeterminação como instrumento regulador dos discursos políticos, caracterizados a partir dos dizeres misógino, homofóbico, ou racista.

Nessa mesma matriz teórica de Meiklejohn e Sunstein, Robert Bork se filia à ideia de liberdade de expressão na autodeterminação coletiva. Contudo, sustenta um posicionamento de que a proteção constitucional, diferentemente das propostas dos demais autores, somente deveria resguardar os discursos cuja essência política seja eminente. Nessa linha, não abrindo a possibilidade de intervenção judicial que tenha como desiderato ampliar a proteção em favor de outros discursos, como expressões artísticas, Bork consigna, ainda, que entre os discursos caracterizados como políticos, não pode haver a possibilidade de “obstrução constitucional a lei tornando crime qualquer tipo de discurso que advogue a derrubada do governo.”¹¹³

¹¹¹ Conferir, entre outros: WALDRON, Jeremy. **The Harm in Hate Speech**, Cambridge, MA and London, England: Harvard University Press, 2012; HOWARD, Joffrey W. Free speech and hate speech. **Annual Review of Political Science**, v. 22, n. 1, p. 93-109, 2019. Disponível em: <https://doi.org/10.1146/annurev-polisci-051517-012343>. Acesso em: 20/12/2021; OMMATI, José Emílio Medauar. **Liberdade de Expressão e o discurso do ódio na Constituição de 1988**. 5. ed. Belo Horizonte: Conhecimento Editora, 2021.

¹¹² FISS, Owen M. **A Ironia da Liberdade de Expressão**. Rio de Janeiro: Renovar, 2005, p. 36.

¹¹³ BORK, Robert H. Neutral Principles and Some First Amendment Problems. **Indiana Law Journal**, v. 47, Fall 1971, Number I, p. 20. Disponível em: <https://heinonline.org/HOL/LandingPage?handle=hein.journals/indana47&div=8&id=&page=>. Acesso em: 28/12/2021.

Para Bork, a classificação dos tipos de discursos em políticos ou apolíticos configura, em sua essência, um meio de fragilizar a intervenção do Poder Judiciário como forma de evitar ou debelar eventuais abusos, restituindo tal poder ao Parlamento. Percebe-se, portanto, que o pensamento de Bork funda-se na ideia de autodeterminação coletiva.

Verifica-se, então, as teorias da verdade e a democrática tendo importantes distinções, contudo, encontrando um eixo comum, qual seja, ambas não protegem integralmente os discursos. A teoria democrática prioriza a proteção do autogoverno e os demais discursos não recebem a mesma atenção pela constituição. Ademais, Meiklejohn não esclarece o que deveria ser classificado como relevante no discurso político. É necessário admitir, portanto, “que essa teoria apresenta limites abstratos imprecisos, para não dizer ambíguos.”¹¹⁴

Nessa linha, Rodrigues Júnior entende como desvantagem dessa teoria o fato de privilegiar o discurso de conteúdo político em detrimento de outras formas de expressão, o que poderia levar a liberdade da expressão à regulação estatal, sendo certo que, enquanto o discurso político seria absolutamente imune, qualquer forma de expressão de conteúdo não-político estaria sujeita a controle e regulação.¹¹⁵

Por fim, registramos que Costa Neto, por sua vez, preconiza os direitos fundamentais não estando livremente à disposição da maioria, daí porque nem mesmo o povo se encontra autorizado a vedar as formas de discurso político protegidas pela liberdade de expressão. Critica, ainda, o fato dessa teoria ser insuficiente quando se pensa em formas de expressão sem conotação política imediata, pois este não contribui para o processo democrático, uma vez que não visa alterar a sociedade em que vivemos¹¹⁶.

1.2.4. A liberdade de expressão como garantia da estabilidade governamental

Essa teoria apresenta a liberdade de expressão como uma das condições essenciais para garantir a estabilidade governamental, posto que um estado se torna mais forte, quando alicerçado na ampla liberdade de expressão, sendo menos sujeito a conflitos sociais do que um estado autoritário, que ao contrário, se funda no base no medo, repressão e perseguição¹¹⁷.

¹¹⁴ LAURENTIIS, Lucas Catib de; THOMAZINI, Fernanda Alonso. Liberdade de expressão: Teorias, Fundamentos e Análise de Casos. **Revista Direito e Práxis**. Rio de Janeiro, v. 11, n. 4, 2020, p. 2260-2301, p. 2274/2275.

¹¹⁵ RODRIGUES JÚNIOR, Álvaro. **Liberdade de expressão e liberdade de informação**: limites e formas de controle. Curitiba: Juruá, 2009, p. 67.

¹¹⁶ COSTA NETO, João. **Liberdade de expressão**: o conflito entre o legislador e o juiz constitucional. São Paulo: Saraiva, 2017, p. 45.

¹¹⁷ MACHADO, Jónatas Eduardo Mendes. **Liberdade de Expressão**: dimensões constitucionais da esfera pública no sistema social. Coimbra: Coimbra Editora, 2002. p. 282.

Nesse sentido, o Estado privilegia a lógica, em vez da força, não sendo saudável para sua estabilidade a supressão ou repressão da expressão, que até pode silenciar ou frear mudanças sociais momentaneamente, sem, contudo, conseguir estabelecer uma relação de lealdade ou unidade com os indivíduos. Desta forma, uma sociedade que prestigia a liberdade de expressão terá mais possibilidades de promover mais tolerância, uma vez que a sociedade tende a aceitar a coexistência de opiniões, ideias e modos de vida diversos¹¹⁸.

A liberdade de expressão produzirá mais estabilidade e menos violência, pelos seguintes motivos: Primeiro, porque as pessoas estão mais propícias a confiar em um governo disposto a ouvir e ter consideração pelos seus argumentos. Ao contrário, se as pessoas visualizam o governo como irracional ou arbitrário, logicamente terão menos confiança nesses governos, bem como taxarão seus líderes como não confiáveis e haverá uma diminuição no respeito e cumprimento das leis. Segundo, se é oportunizado aos indivíduos fazer críticas à política governamental, possibilitando a liberdade de estabelecer suas convicções e persuadir outras pessoas a adotá-las, havendo após esse processo a elaboração de lei, tais indivíduos estarão mais propensos a segui-las e obedecê-las, pois foi assegurado sua participação no processo decisório político¹¹⁹.

Nessa linha, o Estado estará obrigado a controlar as condições em face das quais a liberdade de expressão pode funcionar, garantindo, com isso, um bem-estar geral. Esse controle decorre da responsabilidade de ilidir injustiças que iniciem desordens, como também da manutenção de condições econômicas e sociais favoráveis para o funcionamento da democracia.

Chequer compreende essa teoria da liberdade de expressão envolvendo mais do que uma simples técnica de alcançar o melhor julgamento social, para ele compreende uma visão de sociedade, crença e um meio de vida. É, portanto, prescrita para alcançar uma criação progressiva e uma comunidade intelectualmente potente, estabelecendo um meio de vida que energize à tolerância, o ceticismo, as reações e iniciativas. Desencorajando, por outro lado, a possibilidade de uma sociedade estagnada, conformada, irracional ou tirânica¹²⁰.

¹¹⁸ COSTA NETO, João. **Liberdade de expressão**: o conflito entre o legislador e o juiz constitucional. São Paulo: Saraiva, 2017, p. 42.

¹¹⁹ SCHAUER, Frederick. **Free Speech**: a philosophical enquiry. London: Cambridge University Press, 1984, p. 79.

¹²⁰ CHEQUER, Cláudio Márcio de Carvalho. **A liberdade de expressão como direito fundamental preferencial prima facie** (análise crítica e proposta de revisão ao padrão jurisprudencial brasileiro). 2010. 340 f. Tese de Doutorado - Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2010. p. 43.

1.2.5. Síntese conclusiva

Isoladamente, nenhuma das teorias que argumentam a liberdade de expressão como direito fundamental parece abranger materialmente todos os diferentes tipos de discurso e nesse sentido, entende-se apropriado defender uma natureza multifuncional para a liberdade de expressão como constituinte de um sistema, como aponta Jonatas Machado.

Não se exclui de todo o debate teórico realizado sobre propostas que demonstram as justificativas filosóficas da liberdade de expressão, em sentido amplo, na verdade, verifica-se que todas se retroalimentam, fortalecendo e contribuindo para o enriquecimento dessa liberdade. Em que pese essa percepção de apenas uma teoria apresentar soluções para o problema, sofrendo com isso críticas, a solução está justamente no somatório dessas propostas teóricas, apresentando justificativas teóricas múltiplas para a liberdade de expressão¹²¹ e, consequentemente, os argumentos instrumentais e constitutivos não são excludentes, entre si.

Nesse sentido, Dworkin endossa que “as justificações instrumental e constitutiva da liberdade de expressão não excluem uma à outra,”¹²² e conclui:

“Os dois tipos de justificação, além disso, têm muitos pontos em comum. Nenhum deles atribui um caráter absoluto à liberdade de expressão; ambos admitem que os valores por eles invocados podem ser postos em segundo plano em casos especiais: ao se decidir, por exemplo, até que ponto se devem censurar as informações militares. Não obstante, as duas justificações são essencialmente diferentes, pois a justificação instrumental é mais frágil e mais limitada. É mais frágil porque, como veremos, existem circunstâncias em que as metas estratégicas às quais ela faz apelo parecem exigir uma limitação da liberdade de expressão, e não a proteção desta. É mais limitada porque, ao passo que a justificação constitutiva abrange, em princípio, todos os aspectos da expressão ou do pensamento cuja independência é exigida pela responsabilidade moral, a justificação instrumental, pelo menos em suas versões mais populares, trata principalmente da proteção da expressão política.”¹²³

Apresentadas as finalidades da liberdade de expressão, verifica-se a importância do estudo do limite e alcance desse direito fundamental.

1.3. Liberdade de expressão: um direito fundamental limitado?

¹²¹ MACHADO, Jónatas Eduardo Mendes. **Liberdade de Expressão: dimensões constitucionais da esfera pública no sistema social**. Coimbra: Coimbra Editora, 2002. p. 290.

¹²² DWORKIN, Ronald. **O Direito da Liberdade: a leitura mora da constituição norte-americana**. Tradução Marcelo Brandão Cipolla. 2ª ed. São Paulo: Editora WMF Martins Fontes, 2019, p. 321.

¹²³ DWORKIN, Ronald. **O Direito da Liberdade: a leitura mora da constituição norte-americana**. Tradução Marcelo Brandão Cipolla. 2ª ed. São Paulo: Editora WMF Martins Fontes, 2019, p. 321.

Prevalece, doutrinariamente, a ideia de que o direito fundamental absoluto ou ilimitado¹²⁴, que não admita restrições, está superada.¹²⁵ Durante um tempo, teve-se a ideia de um direito fundamental absoluto, possuidor de bases na fundamentação jusnaturalista, decorrente do conceito racionalista do Direito Natural do pensamento de John Locke, sob a perspectiva de que qualquer violação ou restrição aos direitos fundamentais resultava, em certa medida, contra a natureza. Ingo Sarlet indica a contribuição doutrinária de Locke, tendo em vista que fora “o primeiro a reconhecer aos direitos naturais e inalienáveis do homem, como a vida, liberdade e propriedade, uma eficácia oponível, inclusive, aos detentores do poder”,¹²⁶ baseado no contrato social, sendo que apenas “os cidadãos teriam o direito de resistência, reconhecidos como verdadeiros sujeitos, e não meros objetos do governo.”¹²⁷

Tem-se, assim a predominância da concepção de direitos fundamentais como limitados ou restringíveis, decorrente principalmente do caráter universal ou geral dos direitos fundamentais, bem como da necessidade desses direitos coexistirem entre si, ou com outros direitos fundamentais protegidos constitucionalmente.

Kanayama descreve que os direitos fundamentais só podem ser limitados se estiverem constitucionalmente previstos, servindo até como inibidores da discricionariedade do aplicador da lei, contudo, para casos não expressamente previstos no texto constitucional, é necessária uma análise mais detalhada. Com isso, aponta sobre a necessidade das restrições aos direitos fundamentais serem cuidadosamente limitadas, devendo ter cautela para não chegar ao ponto de impedir restrição de direitos, ao ponto de prejudicar a existência de outros bens constitucionais, ou da própria sociedade.¹²⁸ E arremata:

¹²⁴ A noção de direito fundamental absoluto, predominou até o Século XVIII. Sobre o tema consultar: SACAMAZANO, Joaquin Brage. **Los Limites a los Derechos Fundamentales**. Madrid, 2004.

¹²⁵ Reconhece-se a existência de discussão sobre o direito da não tortura e a proibição de censura poderiam ser caracterizados como direitos absolutos. Contudo, tem prevalecido o entendimento que esses direitos seriam, na verdade, como limites dos direitos fundamentais à integridade física e à liberdade de expressão, direitos esses, que aceitam restrições. Nesse sentido, Sarlet e Robl Filho aduzem, “ainda que a liberdade de expressão, para sua necessária proteção, exija uma interpretação ampla de censura, importa distinguir a censura de outras modalidades de restrição (que poderão, a depender do caso, ser constitucionalmente justificadas), o que se impõe até mesmo para preservar as peculiaridades de cada modalidade da liberdade de expressão e até mesmo para não atribuir caráter absoluto à liberdade de expressão.” (ROBL FILHO, Ilton; SARLET, Ingo W. Estado democrático de direito e os limites da liberdade de expressão na Constituição Federal de 1988, com destaque para o problema da sua colisão com outros direitos fundamentais, em especial, com os direitos de personalidade. **Constituição, Economia e Desenvolvimento: Revista da Academia Brasileira de Direito Constitucional**. Curitiba, 2016, v. 8, n. 14, Jan.-Jun. p. 112-142. p. 124).

¹²⁶ SARLET, Ingo Wolfgang. **A eficácia dos direitos fundamentais**: uma teoria geral dos direitos fundamentais na perspectiva constitucional. 10 ed. 2. Tir. Porto Alegre: Livraria do Advogado Editora, 2010. p. 40.

¹²⁷ SARLET, Ingo Wolfgang. **A eficácia dos direitos fundamentais**: uma teoria geral dos direitos fundamentais na perspectiva constitucional. 10 ed. 2. Tir. Porto Alegre: Livraria do Advogado Editora, 2010. p. 40.

¹²⁸ KANAYAMA, Rodrigo Luís. **Limites às restrições aos direitos fundamentais na atividade de investigação do Poder Legislativo**. 2007. 218 p. Dissertação (Mestrado em Direito do Estado) – Setor de Ciências Jurídicas, Universidade Federal do Paraná, Curitiba, 2007. p. 89.

Defende-se a possibilidade da restrição dos direitos fundamentais. Porém, não se concorda com sua disponibilidade, ou seja, é defeso que a maioria ou a unanimidade possa afastar a aplicação, efetividade ou vigência de um direito fundamental.¹²⁹

Nesse ponto, o principal questionamento quanto à limitação ou à restrição do direito fundamental decorre do estabelecimento de ser ou não intrínseco ao próprio direito fundamental ou se é decorrente de fatores externos. A importância dessa discussão cinge em razão da consequência que pode gerar de sua resposta.

Nesse debate de limitação de direito fundamental, Robert Alexy¹³⁰ apresenta duas teorias, a saber: a teoria interna (ou concepção estrita) dos limites dos Direitos Fundamentais e a teoria externa (ou concepção ampla) dos limites dos Direitos Fundamentais¹³¹.

A teoria interna: segundo essa teoria, a própria Constituição já determina previamente e definitivamente o conteúdo atinente a todos os direitos fundamentais, bem como já fixa todos os seus limites e, sendo o conteúdo dos direitos fundamentais fixado na Carta Magna, eventuais atos legislativos e medidas empreendidas pelos demais poderes públicos estariam impedidos de restringir tais direitos, pois, não haveria margem capaz de restrição externa, posto que a delimitação detalhada do conteúdo do direito fundamental já fora realizada pela própria Constituição¹³².

Nessa linha, verifica-se a importância dessa teoria, também, no campo da interpretação judicial, pois, se o direito fundamental já está preestabelecido de forma definitiva no texto constitucional, restará impossível haver uma situação na qual se possa constatar um conflito entre direitos fundamentais.

Para Carlos Bernal Pulido, o papel do intérprete é desvendar a demarcação de cada um dos direitos fundamentais e vencida essa etapa, vislumbra-se a aplicação única de cada direito fundamental, para cada situação concreta enfrentada¹³³.

Assim, Chequer explica que a “normatividade dos direitos fundamentais não se pode esgotar no texto da norma”¹³⁴, para ele “as disposições normativas não podem ser aplicadas

¹²⁹ KANAYAMA, Rodrigo Luís. **Limites às restrições aos direitos fundamentais na atividade de investigação do Poder Legislativo**. 2007. 218 p. Dissertação (Mestrado em Direito do Estado) – Setor de Ciências Jurídicas, Universidade Federal do Paraná, Curitiba, 2007. p. 89.

¹³⁰ ALEXY, Robert. **Teoria dos Direitos Fundamentais**. Tradução Virgílio Afonso da Silva. 2 ed. São Paulo: Malheiros Editores, 2011. p. 277 e seguintes.

¹³¹ No Brasil, conferir a respeito a completa obra de PEREIRA, Jane Reis Gonçalves. **Interpretação Constitucional e Direitos Fundamentais**. Rio de Janeiro: Renovar, 2006. p. 131 e ss.

¹³² PULIDO, Carlos Bernal. **El principio de proporcionalidad y los derechos fundamentales**: el principio de proporcionalidad como critério para determinar el contenido de los derechos fundamentales vinculante para el legislador. 3ª ed. Madrid: Centro de Estudios Políticos y Constitucionales. 2007. p. 453 e ss.

¹³³ PULIDO, Carlos Bernal. **El principio de proporcionalidad y los derechos fundamentales**: el principio de proporcionalidad como critério para determinar el contenido de los derechos fundamentales vinculante para el legislador. 3ª ed. Madrid: Centro de Estudios Políticos y Constitucionales. 2007. p. 453 e ss.

como prescrições fechadas em si mesmas, devendo, portanto, serem concretizadas conforme o caso concreto”¹³⁵. Nesse sentido, fica demonstrada a não coadunação dessa teoria com a teoria da ponderação de princípios, a qual aponta como diretriz a possibilidade de dois direitos fundamentais entrarem em colisão de acordo com o caso concreto. A teoria interna, por sua vez, entende que a ponderação não segue uma racionalidade metodológica, caracterizando como um critério subjetivo e bastante irracional¹³⁶.

Para a teoria interna, não existe distinção entre o direito e suas restrições, mas somente uma, é dizer, o direito com um determinado conteúdo previamente fixado de forma definitiva pelo constituinte. Nota-se, com isso que para essa teoria o direito fundamental possui um conteúdo bastante reduzido, levando-o a ter muitas situações contempladas pelo direito fundamental da liberdade de expressão.

Nessa perspectiva, Alexy esclarece que o conceito de restrição é substituído pelo conceito de limite, entendendo-se assim, que eventuais dúvidas acerca dos limites não são dúvidas sobre quão extensa pode ser sua restrição, mas uma dúvida sobre o seu conteúdo. Para o autor, quando, eventualmente, se fala em restrição, em vez de limites, refere-se a restrições imanentes. Os direitos fundamentais são concebidos como posições definitivas, não havendo possibilidade de restrição da disponibilidade de direitos fundamentais, todavia, os bens constitucionalmente protegidos podem.¹³⁷

Acrescente-se a essa ideia, o pensamento de Konrad Hesse, segundo ele, os direitos fundamentais são liberdades jurídicas e, como tais, são determinadas materialmente, ou seja, limitadas. Essa limitação seria, então, a determinação de seu alcance material. Assim, como são fundamentadas pela Constituição só nela seus limites podem encontrar base, admitindo, excepcionalmente, a possibilidade de uma limitação não escrita. Ademais, podem ser limitados por outras normas jurídicas, logo, o seu conteúdo é fixado a partir da conexão entre a norma de direito fundamental e a norma de limitação, contudo, a atuação do legislador

¹³⁴ CHEQUER, Cláudio Márcio de Carvalho. **A liberdade de expressão como direito fundamental preferencial prima facie** (análise crítica e proposta de revisão ao padrão jurisprudencial brasileiro). 2010. 340 p. Tese de Doutorado - Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2010. p. 55.

¹³⁵ CHEQUER, Cláudio Márcio de Carvalho. **A liberdade de expressão como direito fundamental preferencial prima facie** (análise crítica e proposta de revisão ao padrão jurisprudencial brasileiro). 2010. 340 p. Tese de Doutorado - Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2010. p. 55.

¹³⁶ CHEQUER, Cláudio Márcio de Carvalho. **A liberdade de expressão como direito fundamental preferencial prima facie** (análise crítica e proposta de revisão ao padrão jurisprudencial brasileiro). 2010. 340 p. Tese de Doutorado - Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2010. p. 55.

¹³⁷ ALEXY, Robert. **Teoria dos Direitos Fundamentais**. Tradução Virgílio Afonso da Silva. 2 ed. São Paulo: Malheiros Editores, 2011. p. 277 e seguintes.

ordinário pode, apenas, declarar os limites já traçados pela Constituição ou determinar os limites da garantia, quando a Constituição autoriza a reserva legal.¹³⁸

Interessante a ideia de Hesse quanto à interpretação de limitações constitucionais ou limites com base em uma reserva legal, segundo ele, deve-se chegar a uma eficácia ótima, não se privando a garantia jurídico-fundamental mais do que o necessário, daí o motivo de seu entendimento sobre como a limitação deve ser adequada e necessária para produzir a proteção do bem jurídico e proporcional como peso e significado do direito fundamental¹³⁹.

A teoria externa: na concepção dessa teoria, exigem-se duas fases de raciocínio para a identificação da proteção de um direito fundamental. Primeiro, seleciona o conceito inicialmente protegido de um direito fundamental. Nessa fase, ao contrário da teoria interna, o raio de proteção do direito fundamental é bem mais amplo. Nela será realizada a harmonização do conteúdo desse direito fundamental, selecionando-se a *prima facie* com os demais direitos fundamentais e bens constitucionalmente protegidos. Após essa fase, encontra-se o conteúdo definitivo do direito fundamental, ou seja, a constatação das hipóteses de que esse direito deve ser decididamente protegido por um direito fundamental.¹⁴⁰

Com efeito, o fato de chegar ao conteúdo do direito fundamental definitivo, após o aludido processo, deixa claro a possibilidade de esse direito (inicialmente determinado) pode estar sujeito a restrições decorrentes de fatores externos.

A teoria externa, com efeito, entende que toda supressão a um direito fundamental é caracterizada como uma limitação externa. Podendo essa limitação ser incorporada pelo ordenamento jurídico, contudo, há necessidade de previsão constitucional específica, bem como de lei limitadora do direito fundamental contendo disposições genéricas, além de aplicar o princípio da proporcionalidade.

Extrai-se, portanto, que a teoria em estudo elege o princípio da proporcionalidade como método para delinear o conteúdo definitivo dos direitos fundamentais, utilizando-o como forma de revelar o argumento definitivo de determinado direito fundamental perante cada nuance do caso concreto.¹⁴¹

¹³⁸ HESSE, Konrad. **Elementos de Direito Constitucionais da República Federal da Alemanha**. Tradução Luis Afonso Heck. Porto Alegre: Sergio Antônio Fabris Editor, 1998, p. 250/255

¹³⁹ HESSE, Konrad. **Elementos de Direito Constitucionais da República Federal da Alemanha**. Tradução Luis Afonso Heck. Porto Alegre: Sergio Antônio Fabris Editor, 1998, p. 255/256

¹⁴⁰ BOROWSKI, Martin. La restricción de los derechos fundamentales. **Revista Española de Derecho Constitucional**. Ano 20, n. 59, p. 29-56, 2000. p. 31/32. Disponível em: <https://dialnet.unirioja.es/servlet/articulo?codigo=79684>. Acesso em: 10/01/2022.

¹⁴¹ Nesse sentido conferir MUÑOZ, Mauricio Maldonado. El principio de razoabilidad y su aplicación al estudio de validez de las normas jurídicas. **Ius Humani. Revista de Derecho**. v. 3, p. 139-174, 2012/2013. p. 155. Disponível em: <https://www.iushumani.org/index.php/iushumani/article/view/40>. Acesso em 15/01/2022.

Essa corrente é amplamente aceita, sendo alguns de seus principais expoentes, dentre outros, Robert Alexy, Martin Borowski, Manuel Medina Guerreiro, Joaquim Brage Camazano, Pedro Serna e Fernando Toller. No Brasil, a teoria externa é aceita majoritariamente, sendo a ponderação de princípios um dos métodos de orientação da interpretação para solução de conflitos entre direitos fundamentais.¹⁴²

Alexy, partindo da ideia de “restrição a direitos fundamentais”, questiona a necessidade de identificar qual o teor dessas restrições, denotando-as da seguinte forma, “restringíveis são os bens protegidos por direitos fundamentais (liberdades/situações/posições de direito ordinário) e as posições *prima facie* garantidas por princípios de direitos fundamentais.”¹⁴³ Além de os princípios de direitos fundamentais, ao contrário, demandarem uma proteção mais abrangente possível “da liberdade de ação, da integridade física ou da competência para alienar a propriedade”¹⁴⁴. Nesse raciocínio, Alexy apresenta uma resposta à indagação sobre tais restrições: “restrições a direitos fundamentais são normas que restringem uma posição *prima facie* de direito fundamental.”¹⁴⁵ Deduz, portanto, serem as restrições a direitos fundamentais normas as quais poderão restringir o direito fundamental, se for compatível com a Constituição. A norma restritiva pode ser uma regra ou um princípio.

Alexy, prosseguindo em neste raciocínio, demonstra que uma *regra* pode ser entendida como restrição a um direito fundamental, conquanto, ao ser editada, apodere-se de uma liberdade *prima facie*, ou mesmo de um direito fundamental *prima facie*, se dessa substituição surja uma “não-liberdade definitiva ou um não-direito definitivo de igual conteúdo”¹⁴⁶. Para ilustrar seu raciocínio exemplifica, ainda, que enquanto não houver uma determinação aos motociclistas para usar capacete, em razão do princípio da liberdade geral de ação (uma liberdade fundamental *prima facie*), eles podem decidir entre usar ou não o capacete. Contudo, se houver edição de regra determinando a obrigatoriedade de utilizar capacete, o titular deste direito fundamental passa para a posição de não-liberdade definitiva. E conclui: “a regra que obriga os motociclistas a usar capacete é uma restrição a um direito

¹⁴² Com essa ideia, Sarmento expõe que: “ao que os juízes e intérpretes fazem de fato nestes casos difíceis – pesam e contrapõem interesses, valores, argumentos – e por indicar a eles um caminho mais racional e controlável para esta empreitada hermenêutica.” (SARMENTO, Daniel. A liberdade de expressão e o problema do “Hate Speech”. In SARMENTO, Daniel. **Livres e iguais**: estudos de Direito Constitucional. Rio de Janeiro: Lumen Juris, 2006. p. 260.

¹⁴³ ALEXY, Robert. **Teoria dos Direitos Fundamentais**. Tradução Virgílio Afonso da Silva. 2 ed. São Paulo: Malheiros Editores, 2011. p. 281.

¹⁴⁴ ALEXY, Robert. **Teoria dos Direitos Fundamentais**. Tradução Virgílio Afonso da Silva. 2 ed. São Paulo: Malheiros Editores, 2011. p. 281.

¹⁴⁵ ALEXY, Robert. **Teoria dos Direitos Fundamentais**. Tradução Virgílio Afonso da Silva. 2 ed. São Paulo: Malheiros Editores, 2011. p. 281.

¹⁴⁶ ALEXY, Robert. **Teoria dos Direitos Fundamentais**. Tradução Virgílio Afonso da Silva. 2 ed. São Paulo: Malheiros Editores, 2011. p. 283.

fundamental, porque em virtude de sua vigência surge, no lugar da liberdade *prima facie*, uma não-liberdade definitiva de igual conteúdo.”¹⁴⁷

Por outro lado, os *princípios* que restringem, entretanto, não podem, por si só, colocar o indivíduo em determinadas posições definitivamente limitadoras, ou seja, uma não-liberdade ou um não-direito de igual conteúdo. Nesse caso é necessário o sopesamento entre o princípio constitucional atingido e o princípio que o restringe.¹⁴⁸ As restrições podem ser classificadas como diretamente constitucionais, ou seja, norma de hierarquia constitucional ou indiretamente constitucional, normas infraconstitucionais decorrentes de normas constitucionais.¹⁴⁹

Quanto às restrições indiretamente constitucionais, ou também denominadas de cláusulas de reserva, Alexy¹⁵⁰ leciona que são aquelas que a Constituição autoriza alguém a estabelecer, elas podem ser explícitas ou implícitas, ou seja, são explícitas quando as disposições constitucionais, ou parte delas, expressamente autorizam intervenções e restrições. Nesse ponto distingue-se reserva simples e qualificada, na primeira a competência para estabelecer restrições é pura e simplesmente garantida, no segundo caso há uma limitação ao conteúdo da restrição.

Nesse particular, entretanto, o principal problema surgido é sua delimitação, principalmente, do ponto de vista material, pois, do ponto de vista formal está adstrito à competência para impor restrições ao procedimento e à forma. Já do ponto de vista material é de se afirmar que o legislador não está subordinado a estabelecer aquilo já existente, todavia, deve ficar limitado não somente “pelas condições expressas nas reservas qualificadas e pela barreira do conteúdo essencial”¹⁵¹, além de respeitar o princípio da proporcionalidade. Assim com essa junção entre competência e sopesamento, evita-se, por um lado, que os direitos fundamentais não fiquem na dependência do legislador, e, com isso, tenha seu conteúdo

¹⁴⁷ ALEXY, Robert. **Teoria dos Direitos Fundamentais**. Tradução Virgílio Afonso da Silva. 2 ed. São Paulo: Malheiros Editores, 2011. p. 283.

¹⁴⁸ ALEXY, Robert. **Teoria dos Direitos Fundamentais**. Tradução Virgílio Afonso da Silva. 2 ed. São Paulo: Malheiros Editores, 2011. p. 283/284.

¹⁴⁹ Restrições a direitos fundamentais, segundo a visão de Alexy, ou “são normas de hierarquia constitucional ou normas infraconstitucionais, que possuem criação com a devida autorização por normas constitucionais. Desse modo, as restrições de hierarquia constitucional são restrições *diretamente constitucionais*, e as restrições infraconstitucionais são restrições *indiretamente constitucionais*”. (ALEXY, Robert. **Teoria dos Direitos Fundamentais**. Tradução Virgílio Afonso da Silva. 2 ed. São Paulo: Malheiros Editores, 2011. p. 286).

¹⁵⁰ ALEXY, Robert. **Teoria dos Direitos Fundamentais**. Tradução Virgílio Afonso da Silva. 2 ed. São Paulo: Malheiros Editores, 2011. p. 291.

¹⁵¹ ALEXY, Robert. **Teoria dos Direitos Fundamentais**. Tradução Virgílio Afonso da Silva. 2 ed. São Paulo: Malheiros Editores, 2011. p. 292.

esvaziado.¹⁵² Entretanto, pode surgir uma submissão muito forte aos direitos fundamentais, e isso ocasione que o legislador, tão somente, estabeleça o resultado de um sopesamento entre princípios relevantes. Alexy resume essa preocupação, pontuando “a legislação sob reserva constitucional nada mais seria que interpretação das restrições já existentes em virtude do texto constitucional.”¹⁵³

No contexto brasileiro, a Constituição Federal de 1988 não estipula um rol de hipóteses de restrição à liberdade de expressão¹⁵⁴, mas é possível verificar algumas restrições explícitas no texto constitucional, como é o caso de propaganda comercial de tabaco, e outras implícitas de alguns preceitos constitucionais, como é o caso de direitos de personalidade.

Nessa linha de pensamento, Carvalho¹⁵⁵ pontua que a liberdade de expressão encontra limites na própria constituição, estando limitada pela proteção assegurada, constitucionalmente, dos direitos de personalidade. Contudo, em face do grau de relevância conferido pela Constituição, estaria a salvo de certas investidas do poder público, sendo estas limitáveis, controláveis e dependentes de permissão constitucional. Aos agentes administrativos não lhes caberia imiscuírem-se na liberdade de expressão, a título de controlar sua legitimidade, até por não caber a eles avaliar subjetivamente qual o bem jurídico deve ponderar. Essa missão seria conferida ao Poder Judiciário, pois possui competência para sopesar os valores constitucionais, estando, ainda, adstrito a impor restrições imprescindíveis para salvaguardar outros direitos sem proteção por meios mais gravosos, prevalecendo a possibilidade de reparação por indenização.¹⁵⁶

Por sua vez, Farias contribui para o debate aduzindo, a despeito do inestimável valor para o indivíduo e para a sociedade, a liberdade de expressão sujeita-se a vários tipos de restrições, todas decorrentes, direta ou indiretamente da constituição, implícitas ou explícitas.

¹⁵² ALEXY, Robert. **Teoria dos Direitos Fundamentais**. Tradução Virgílio Afonso da Silva. 2 ed. São Paulo: Malheiros Editores, 2011. p. 292.

¹⁵³ ALEXY, Robert. **Teoria dos Direitos Fundamentais**. Tradução Virgílio Afonso da Silva. 2 ed. São Paulo: Malheiros Editores, 2011. p. 293.

¹⁵⁴ Sarlet, quanto ao tema de limitação à liberdade de expressão no Brasil, aduz: “Dada a sua relevância para a democracia e o pluralismo político, a liberdade de expressão – pelo menos de acordo com significativa doutrina – assume uma espécie de posição preferencial (*preferred position*), quando da resolução de conflitos com outros princípios constitucionais e direitos fundamentais, o que tem sido, em regra, confirmada pelo STF em especial a partir do julgamento da ADPF n. 130, no qual foi declarada que a Lei de Imprensa editada no período militar não foi recepcionada pela CF”. (SARLET, Ingo Wolfgang. Direitos Fundamentais em espécie. In: SARLET, Ingo Wolfgang; MITIDIERO, Daniel; MARINONI, Luiz Guilherme. **Curso de direito constitucional**. 9 ed. São Paulo: Saraiva Educação, 2020. p. 518)

¹⁵⁵ CARVALHO, Luis Gustavo Grandinetti Castanho de. **Liberdade de informação e o direito difuso à informação verdadeira**. 2º ed. Rio de Janeiro: Renovar, 2000, p. 49-51.

¹⁵⁶ Sobre o direito fundamental à indenização por danos materiais e imateriais causados com base no exercício da liberdade de expressão, conferir: SARLET, Ingo Wolfgang. Direitos Fundamentais em espécie. In: SARLET, Ingo Wolfgang; MITIDIERO, Daniel; MARINONI, Luiz Guilherme. **Curso de direito constitucional**. 9 ed. São Paulo: Saraiva Educação, 2020. p. 524 e seguintes.

Segundo a ótica do autor, dão-se de três maneiras: direitos fundamentais dos demais cidadãos (os direitos de personalidade), bens sociais (proteção à saúde, à segurança pública e à imagem), valores estatais e constitucionais (resguardo da própria Constituição ou do Estado).

Destaca, ainda, como restrições *diretamente* constitucionais a vedação ao anonimato; a inadmissibilidade de invocar a liberdade para se eximir de obrigação legal a todos impostas e recusar-se a cumprir prestação alternativa; a inviolabilidade do direito à honra, à intimidade, à vida privada e à imagem; sigilo das informações imprescindíveis à segurança do Estado e da Sociedade; proibição de monopólio e oligopólio quanto à propriedade dos meios de comunicação; conteúdo da programação das emissoras de rádio e televisão; necessidade de outorga do poder público para a exploração de serviço de radiodifusão sonora e de sons e imagens.

Elenca, ainda, como *indiretamente* constitucionais a necessidade de qualificação profissional para o exercício da comunicação social; a intimidade e o interesse social como exceção à publicidade dos atos processuais; restrição à diversos espetáculos públicos; a restrição quanto à propaganda comercial de tabaco, bebida alcoólica, agrotóxico, medicamentos e terapia; restrição à liberdade de imprensa quando em Estado de Sítio.

Acrescenta, por fim, a possibilidade de *restrições tácitas* formuladas pelo legislador, como a proteção da dignidade da criança e do adolescente, o respeito às leis penais e a observância do princípio da presunção de inocência; bem como pelo Judiciário, tais como o respeito aos direitos de personalidade e o resguardo da independência e imparcialidade do Poder Judiciário.¹⁵⁷

Paulo Gonet Branco e Gilmar Mendes, reconhecem como limites a vedação ao anonimato; a imposição do direito de resposta e indenização por danos morais, materiais e à imagem; a preservação da intimidade, vida privada, honra e imagem; a exigência de qualificação profissional dos que se dedicam aos meios de comunicação; o direito de acesso à informação; a restrição à publicidade de bebidas alcóolicas, tabacos, medicamentos e terapias; respeito aos valores éticos e sócias da pessoa e da família; a proteção da segurança pública; a proteção da criança e da juventude; a vedação ao discurso de ódio e pornografia; e a vedação à informação falsa.¹⁵⁸

Conforme delineado, tem-se que o direito fundamental à liberdade de expressão é um direito propenso à restrição ou à limitação, na mesma esteira que outra liberdade fundamental,

¹⁵⁷ FARIAS, Edmilson. **Liberdade de expressão e comunicação: teoria e proteção constitucional**. Revista dos Tribunais, São Paulo, 2004, p. 241-289.

¹⁵⁸ BRANCO, Paulo Gustavo G. Liberdades. In: MENDES, Gilmar Ferreira; BRANCO, Paulo Gustavo. **Curso de direito constitucional**. 15. ed., São Paulo: Saraiva, 2020. p. 275-276.

ficando, contudo, aberto o debate quanto à extensão da restrição ou limitação. Com isso, resta reconhecer unidade ao se admitir como limites a vedação ao anonimato, os direitos de personalidade, a proteção da infância e adolescência, e mais recentemente o discurso ao ódio.¹⁵⁹

1.4. A liberdade de expressão na era digital

No ano de 1926, durante uma entrevista para John B. Kennedy, o famoso cientista sérvio-americano Nicolas Tesla previu aquilo que seria mais tarde os *smartphones*, ao revelar que, quando a tecnologia sem fio tivesse sua aplicação perfeita, a terra inteira seria como um grande cérebro, possibilitando a comunicação instantânea, independentemente da distância, com a possibilidade de vermos e ouvirmos uns aos outros perfeitamente, como se tivéssemos frente a frente, concluindo: “e os instrumentos através dos quais seremos capazes de fazer isso serão incrivelmente simples comparados com o nosso telefone atual. Um homem poderá carregar um no bolso do colete.”¹⁶⁰

A impressionante previsão tornou-se realidade cerca de 100 anos depois, alinhada com a evolução tecnológica revolucionando a comunicação e a interação social, inclusive, denominando-se revolução digital¹⁶¹, possibilitando ampla participação e interação cultural a qual antes não poderia existir na mesma escala. A era digital torna a produção e distribuição

¹⁵⁹ A respeito do tema Duque e Harff, consignam: “O sistema brasileiro, por sua vez, também condena o discurso de ódio, a exemplo dos julgados colacionados, sendo o caso do HC nº 82.424/RS um marco na jurisprudência dos direitos humanos no âmbito do Supremo Tribunal Federal, firmando-se o entendimento em favor da restrição da liberdade de expressão e em favor da dignidade humana e do princípio da igualdade e não discriminação”. (HARFF, Graziela; DUQUE, Marcelo Schenk. Discurso de ódio nos contextos alemão e brasileiro. **A&C – Revista de Direito Administrativo & Constitucional**, Belo Horizonte, ano 21, n. 84, p. 199-225, abr./jun. 2021. p. 222 Disponível em: <https://www.revistaaec.com/index.php/revistaaec/article/view/1350>. Acesso em: 19/11/2021).

¹⁶⁰ *Apud* CHOG, Celena. “**The inventor that inspired elon musk and larry page predicted smartphones nearly 100 years ago**”. Disponível em: <https://www.businessinsider.in/the-inventor-that-inspired-elon-musk-and-larry-page-predicted-smartphones-nearly-100-years-ago/articleshow/47964129.cms>. Acesso em: 25/11/2021.

¹⁶¹ Balkin conceitua da seguinte forma: “Por ‘revolução digital’, entendo a criação e ampla disponibilidade de tecnologias que tornam mais fácil copiar, modificar, anotar, agrupar, transmitir e distribuir conteúdo, armazenando-o em formato digital. Essas tecnologias também incluem o desenvolvimento de vastas redes de comunicação que conectam todas as partes do mundo com o objetivo de distribuir conteúdo digital”. Tradução livre do original: By the “digital revolution,” I mean the creation and widespread availability of technologies that make it easy to copy, modify, annotate, collate, transmit, and distribute content by storing it in digital form. These technologies also include the development of vast communication networks that connect every part of the world for the purpose of distributing digital content”. (BALKIN, Jack M., *Digital Speech and Democratic Culture: A Theory of Freedom of Expression for the Information Society*, **New York University Law Review**, v. 79, p. 1–55, 2004. p. 06. Disponível em: https://openyls.law.yale.edu/bitstream/handle/20.500.13051/1712/Digital_Speech_and_Democratic_Culture_A_Theory_of_Freedom_of_Expression.pdf?sequence=2. Acesso em 04/11/2021).

de informação uma fonte importante de riqueza.¹⁶² Sobretudo, com a versão Web 2.0, com a expansão e abertura da *internet* aos civis¹⁶³, gerando uma nova sociedade em redes¹⁶⁴, contribuindo para a migração de relevantes necessidades da vida aos meios digitais, modificando comportamentos e costumes, além de instituir um novo espaço de interação pública. Lógico, que tudo isso se deu, não somente, por conta da criação dos *smartphones*, mas a visão futurista de Tesla, a qual mais tarde veio a se concretizar, contribuiu para a disseminação, em massa, de um novo mundo digital.¹⁶⁵

Essa inovação e a nova forma de pensar possibilitaram o surgimento de grandes empresas tecnológicas, tais como *Google (Alphabet)*, *Amazon*, *Facebook*, *Apple* e *Microsoft*, conhecidas como as *big five*¹⁶⁶ e as mídias sociais, tornando-se parte do cotidiano da sociedade, influenciando novas formas de manifestação de pensamento, opinião e conhecimento, aumentando, sobremaneira, a troca de informações e debate em tempo real.

Nessa perspectiva, a liberdade de expressão *offline* passa a ser, também *online*, uma forma de digitalização da fala ou discurso, trazendo consigo uma elevação de sua proeminência no constitucionalismo moderno. Além de debates sobre a colisão de direitos nesse ambiente digital, ameaças a direitos fundamentais, como a liberdades de expressão, existência de regulação adequada, participação de atores públicos e privados.

1.4.1. A evolução do discurso *online*

É possível comparar os avanços da era digital com o surgimento da imprensa escrita, ao passo que ambas representaram um divisor na história da utilização e difusão de

¹⁶² BALKIN, Jack M., Digital Speech and Democratic Culture: A Theory of Freedom of Expression for the Information Society, *New York University Law Review*, v. 79, p. 1–55, 2004. p. 02. Disponível em: https://openyls.law.yale.edu/bitstream/handle/20.500.13051/1712/Digital_Speech_and_Democratic_Culture_A_Theory_of_Freedom_of_Expression.pdf?sequence=2. Acesso em 04/11/2021).

¹⁶³ “Em 1994, os blackbones da internet foram privatizados e surgiram os primeiros provedores comerciais. Com isso, graças aos protocolos livres e à arquitetura em múltiplas camadas e distribuída, qualquer pessoa que tivesse um computador podia se conectar à rede mundial de computadores [...]”. (VALENTE, Mariana Giorgetti. A liberdade de expressão na *internet*: da utopia às eras das plataformas. In. FARIA, José Eduardo [Org.]. **A liberdade de expressão e as novas mídias**. 1ª ed. São Paulo: Perspectiva, 2020. p. 25-26)

¹⁶⁴ “No final do século XX vivemos um desses raros intervalos na história. Um intervalo cuja característica é a transformação de nossa ‘cultura material’ pelos mecanismos de um novo paradigma tecnológico que se organiza em torno da tecnologia da informação”. (CASTELLS, Manuel. **A sociedade em rede**. Trad. Roneide Vanâncio Majer. 6ª ed. – A era da informação: economia, sociedade e cultura, v. 1. São Paulo: Paz e Terra, 1992).

¹⁶⁵ MANTOVANI, Camila Maciel C. Alves. **Info-entretenimento na telefonia celular**: informação, mobilidade e interação social em um novo espaço de fluxos. Belo Horizonte, 2006. 152 f. Dissertação (Mestrado em Ciência da Informação) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais.

¹⁶⁶ MOORE, Martin, TAMBINI, Damian. **Digital dominance: the power of Google, Amazon, Facebook, and Apple**. New York. NY: Oxford University Press, 2018. Disponível em: <http://14.139.206.50:8080/jspui/bitstream/1/7241/1/Digital%20Dominance.pdf>. Acesso em 27/11/2021.

liberdades. E, em especial, a facilidade na produção, divulgação e troca de informações, ultrapassando obstáculos físicos e temporais, proporcionando o surgimento de novos discursos, além de tornar-se acessível ao debate, incentivando a participação e engajamento de mais pessoas,

Contudo, com a migração do debate presencial para a *internet*, proporcionou a facilidade de manifestar opiniões e um crescimento enorme de informações, oferecendo um risco à liberdade de se expressar, com a possibilidade dessa liberdade ser controlada, mitigada ou condicionada, seja por ações do Estado ou por instituições privadas.¹⁶⁷

Especialmente, nos anos 90, havia a sensação de, no mundo da *internet*, todos usufruírem de uma liberdade plena, sem interferências externas, sobretudo do Estado, em regular as ações e atividade no ambiente *online*, como aduz Valente “um período em que utopias dominaram as formas de se pensar sobre a *internet*”¹⁶⁸. Um dos personagens mais importantes é o ativista californiano John Perry Barlow o qual durante o Fórum de Econômico Mundial, em Davos, no ano de 1996, lançou a Declaração da Independência do Ciberespaço (*Declaration on the Independence of Cyberspace*), “aproveitando a palavra ciberespaço para codificar um novo domínio de funcionamento, liberdade e responsabilidades”.¹⁶⁹

Na referida "Declaração de Independência do Ciberespaço",¹⁷⁰ Barlow afirma o espaço digital sendo um novo mundo, separado do atômico, onde "os conceitos jurídicos de propriedade, expressão, identidade, movimento e contexto não se aplicam."¹⁷¹ Este mundo virtual e sem uma localização física, também foi apoiado por Johnson e Post,¹⁷² os quais consideram um direito descentralizado e emergente, resultante da ação privada consuetudinária ou coletiva. Ou seja, essas ideias são baseadas em uma abordagem de baixo para cima: em vez de depender do poder legislativo tradicional para definir as regras do

¹⁶⁷ ASH, Timothy Garton. **Free Speech: Ten Principles for a Connected World**. London: Atlantic Book, 2016. p. 20.

¹⁶⁸ VALENTE, Mariana Giorgetti. A liberdade de expressão na *internet*: da utopia às eras das plataformas. In: FARIA, José Eduardo [Org.]. **A liberdade de expressão e as novas mídias**. 1ª ed. São Paulo: Perspectiva, 2020. p. 27.

¹⁶⁹ SILVESTRE, Ricardo. **Liberdade de expressão política na era digital: ameaças e soluções**. 2021. 118 p. Dissertação de Mestrado - Universidade Lusófona de Humanidades e Tecnologias - Faculdade de Ciências Sociais, Educação e Administração, Lisboa, 2021. p. 41.

¹⁷⁰ BARLOW, John P. **Uma Declaração de Independência do Ciberespaço**, Electronic Frontier Found. (8 de fevereiro de 1996), Disponível: www.eff.org/cyberspace-independence. Acesso em: 29/10/2021.

¹⁷¹ BARLOW, John P. **Uma Declaração de Independência do Ciberespaço**, Electronic Frontier Found. (8 de fevereiro de 1996), Disponível: www.eff.org/cyberspace-independence. Acesso em: 29/10/2021.

¹⁷² Conferir: JONHSON, David R. POST, David, Law and Borders: The Rise of Law in Cyberspace. **First Monday**, v. 48 n. 1, 1996. Disponível em: <https://firstmonday.org/ojs/index.php/fm/article/view/468/824>. Acesso em: 20/10/2021.

ciberespaço, toda comunidade digital seria capaz de participar da criação das novas regras que governam seu mundo digital.¹⁷³

Porém, verificou-se a total ausência do Estado no mundo digital apresentando riscos aos direitos fundamentais dos usuários, não vingando a ideia de emancipação total dos cidadãos, agora digitais, às demandas e interferência dos governos. Nesse sentido, afirma Ash: “a *Internet* nunca foi, ou será, totalmente independente da influência de governos, empresas e outros poderes terrestres”¹⁷⁴.

Com a perspectiva de manter esse ambiente *online* o mais próximo possível do que idealizou Barlow, perante fortes tendências de controle e gerência desses espaços por grandes grupos comunicacionais, desenvolveu-se novas políticas de uso e controle defendidas, sobretudo, por organizações da sociedade civil e atores privados, surgindo as principais ações, como a “limitação da responsabilidade de intermediários por conteúdo de terceiros”¹⁷⁵, dentre outras.

Com o avanço da *internet*, possibilitando seu uso para diversas finalidades, notou-se uma evolução maciça nos investimentos nessa área, conforme verificou a alteração do *Silicon Valley*, que evoluiu de pequenos projetos tecnológicos, para desenvolvimento de grandes empresas multinacionais e transnacionais, com capacidade de capital admirável, explorando o novo nicho tecnológico, desempenhando atividades de grande interferência na sociedade com ações que aparentam estruturas governamentais.

O crescimento de empresas privadas no segmento digital, além da grande procura e interesse/necessidades, obteve sucesso por não haver nos Estados Unidos, políticas de regulação dessa atividade, permitindo a cada empresa empregar sua forma de desenvolver produtos, sem nenhuma normatização delimitadora de seu modelo de desenvolvimento de produtos, ou de obtenção de investimentos para seus projetos, seja interno ou externo, permitindo assim, todo o desenvolvimento digital nas mãos da iniciativa privada, diferentemente de outros produtos/serviços do mercado tradicional, adotando um modelo de intervenção estatal mínima.

Dentre os grandes projetos desenvolvidos, merece destaque a disseminação “viral” das denominadas “redes sociais”, as quais, nos EUA, desde 1996, recebem tratamento de isenção

¹⁷³ DE GREGORIO, Giovanni. The rise of digital constitutionalism in the European Union, **International Journal of Constitutional Law**, v. 19, Issue 1, p. 41-70. 2021. p. 44 Disponível em: <https://academic.oup.com/icon/article/19/1/41/6224442>. Acesso em 26/08/2021.

¹⁷⁴ ASH, Timothy Garton. **Free Speech: Ten Principles for a Connected World**. London: Atlantic Book, 2016. p. 22.

¹⁷⁵ VALENTE, Mariana Giorgetti. A liberdade de expressão na *internet*: da utopia às eras das plataformas. In. FARIA, José Eduardo [Org.]. **A liberdade de expressão e as novas mídias**. 1ª ed. São Paulo: Perspectiva, 2020. p. 29.

de responsabilidade sobre o conteúdo produzido nas plataformas, conforme previsão do Código § 230 - Proteção para bloqueio privado e triagem de material ofensivo.¹⁷⁶ Nessa linha de desobrigação, as empresas de mídia social defendem o direito de fazerem *jus* a este tratamento, pois suas atividades não as caracterizam como criadores de conteúdo informacional, tratando-se somente de palco de vivência digital, com fluxo de opiniões e posicionamentos.

Verifica-se essa condição ser primordial para a grande ampliação dessas plataformas de debate, permitindo maior liberdade na participação, sem o receio de eventual responsabilização pelo Estado, por outro lado, é um campo fértil para expansão da polarização política, discurso de ódio, racismo etc.

Esse novo cenário, originário da *internet* e das novas tecnologias, influenciou na alteração das dimensões da luta pelo poder, bem como de seus participantes mais importantes, como as eleições, negócios, política, direito, regulação e o rápido desenvolvimento de tecnologias de comunicação.¹⁷⁷

Nessa linha, Lawrence Lessig sistematiza quatro diferentes tipos de restrições, agindo em qualquer ponto do sistema de informação global: a) o direito; b) o mercado; c) as normas e d) a arquitetura da *internet*. De acordo com o autor, “o código é a lei”, significando que o “software e o hardware (isto é, o ‘código’ do espaço cibernético) que fazem do ciberespaço o que ele é, também regulam como ele é.”¹⁷⁸ Ash, de forma didática, utiliza uma metáfora para descrever a arena onde ocorre a luta pelo poder por meio do ciberespaço: governos são os cachorros, as companhias são os gatos e nós somos os ratos.¹⁷⁹

1.4.2. Características do novo discurso *online*

Na Cúpula Mundial sobre a Sociedade da Informação, Lessig destacou as principais potencialidades oferecidas pelo ambiente digital: “Pela primeira vez em um milênio, temos

¹⁷⁶ Conferir: 47 U.S. Code § 230 - Protection for private blocking and screening of offensive material. Disponível em: <https://www.law.cornell.edu/uscode/text/47/230>.

¹⁷⁷ VALENTE, Mariana Giorgetti. A liberdade de expressão na *internet*: da utopia às eras das plataformas. In. FARIA, José Eduardo [Org.]. **A liberdade de expressão e as novas mídias**. 1ª ed. São Paulo: Perspectiva, 2020. p. 45.

¹⁷⁸ A respeito, verificar: LESSIG, Lawrence. **Code is law: on liberty in cyberspace**. Disponível em: <https://www.havardmagazine.com/>. Acesso em: 04/12/2021.

¹⁷⁹ ASH, Timothy Garton. **Free Speech: Ten Principles for a Connected World**. London: Atlantic Book, 2016. p. 35.

uma tecnologia para equalizar a oportunidade que as pessoas têm de acessar e participar da construção de conhecimento e cultura, independentemente de sua localização geográfica.”¹⁸⁰

A era digital fornece uma infraestrutura tecnológica que expande as possibilidades de participação individual no crescimento e difusão da cultura e, assim, aumenta, na visão de Balkin, as possibilidades de realização de uma cultura verdadeiramente democrática. Contudo, as mesmas tecnologias também podem produzir novos métodos de controle limitadores da participação cultural democrática. Desta forma, exsurge a necessidade dos valores da liberdade de expressão – interatividade, participação em massa e a capacidade de modificar e transformar a cultura – serem protegidos por meio do design tecnológico e da regulamentação administrativa e legislativa da tecnologia, bem como por meio do método mais tradicional de criação judicial e reconhecimento dos direitos constitucionalmente protegidos.¹⁸¹

É inegável um dos avanços oferecido às pessoas pelas tecnologias digitais ser a oportunidade de praticar e participar da criação e desenvolvimento da cultura pública e tais mudanças tecnológicas em como a fala é transmitida e quem participa influenciam em uma nova abordagem da liberdade de expressão¹⁸².

A *internet*, e todos os avanços que representa, permitiu novas perspectivas ao princípio da liberdade de expressão, não que seja totalmente novidade, mas introduz e salienta contornos inéditos. O discurso na *internet* tem maior abrangência de assuntos, modos de expressão, com inclusão de temas sérios, efêmeros, profundos, frívolos e, por assim dizer, reflete gostos populares de todas as classes sociais. Com isso, as pessoas comuns passaram a demonstrar sua enorme criatividade, a qual, até então, não havia chance de expressar, representando uma nova dimensão do poder serem produtores ativos, em vez de meros recipientes passivos de seu mundo cultural¹⁸³.

¹⁸⁰ LESSIG, Lawrence. Uma Sociedade da Informação: Livre ou Feudal. **Cúpula Mundial sobre a Informação Society** (WSIS). 2004. Disponível em: <http://www.itu.int/wsis/docs/pc2/visionaries/lessig.pdf>. Acesso em: 04/12/2021.

¹⁸¹ BALKIN, Jack M. Digital Speech and Democratic Culture: A Theory of Freedom of Expression for the Information Society, **New York University Law Review**, v. 79, p. 1–55, 2004. p. 05. Disponível em: https://openyls.law.yale.edu/bitstream/handle/20.500.13051/1712/Digital_Speech_and_Democratic_Culture_A_Theory_of_Freedom_of_Expression.pdf?sequence=2. Acesso em 04/11/2021.

¹⁸² BALKIN, Jack M. Digital Speech and Democratic Culture: A Theory of Freedom of Expression for the Information Society, **New York University Law Review**, v. 79, p. 1–55, 2004. p. 31. Disponível em: https://openyls.law.yale.edu/bitstream/handle/20.500.13051/1712/Digital_Speech_and_Democratic_Culture_A_Theory_of_Freedom_of_Expression.pdf?sequence=2. Acesso em 04/11/2021.

¹⁸³ BALKIN, Jack M. Digital Speech and Democratic Culture: A Theory of Freedom of Expression for the Information Society, **New York University Law Review**, v. 79, p. 1–55, 2004. p. 31. Disponível em: https://openyls.law.yale.edu/bitstream/handle/20.500.13051/1712/Digital_Speech_and_Democratic_Culture_A_Theory_of_Freedom_of_Expression.pdf?sequence=2. Acesso em 04/11/2021.

O protagonismo, agora compartilhado, imiscui os papéis de leitor e escritor, produtor e consumidor de conteúdo e informações, permitindo ao cidadão possuir postura ativa e ser ouvido e visto, saindo da zona de penumbra, dando voz a todos, proporcionando a chance de utilizar a tecnologia para participar da cultura, interagir, criar, construir, encaminhar, retirar do velho e produzir o novo, além de falar o que quiserem falar, seja política, questões públicas, ou cultura popular.¹⁸⁴

Nesse panorama, é importante a reflexão realizada por Balkin, ao analisar como o entendimento sobre a Primeira Emenda mudará com a era digital: “enfrentamos uma transição de enorme ironia”¹⁸⁵, se referindo quanto ao momento em que a sociedade está com a vida econômica e social cada vez mais influenciada pela tecnologia da informação e pelos fluxos de informação, resultando no fato de a Primeira Emenda parecer cada vez mais irrelevante para os principais entraves pela liberdade de expressão do futuro. Ou seja, os valores-chave da liberdade de expressão¹⁸⁶, logicamente persistem, tanto no mundo off-line como no mundo *online*, possuindo a mesma importância nos blogs, mecanismos de buscas e redes sociais, como sempre possuiu na mídia impressa, panfletos etc. A verdadeira alteração ocorreu no “contexto tecnológico em que tentamos realizar esses valores.”¹⁸⁷

Nesse contexto, Balkin vaticina que as decisões mais importantes, que afetam o futuro da liberdade de expressão, não ocorrerão no âmbito do direito constitucional, mas, sobretudo nas decisões sobre design tecnológico, regulamentos legislativos e administrativos, formação de novos modelos de negócios e nas atividades coletivas dos usuários finais.

Como visto, a inovação digital trouxe avanços à liberdade de expressão, levando a tecer novos contornos quanto à sua utilização nesse novo espaço, como afirma Balkin “a ascensão das redes digitais como tecnologia dominante para a fala em nossa época transforma

¹⁸⁴ BALKIN, Jack M. Digital Speech and Democratic Culture: A Theory of Freedom of Expression for the Information Society, **New York University Law Review**, v. 79, p. 1–55, 2004. p. 42. Disponível em: https://openyls.law.yale.edu/bitstream/handle/20.500.13051/1712/Digital_Speech_and_Democratic_Culture_A_Theory_of_Freedom_of_Expression.pdf?sequence=2. Acesso em 04/11/2021.

¹⁸⁵ BALKIN, Jack M. The Future of Free Expression in a Digital Age, **Pepperdine Law Review**, v. 36, Iss. 2, 2009. p. 427. Disponível em: <https://digitalcommons.pepperdine.edu/plr/vol36/iss2/9/>. Acesso em: 04/11/2021.

¹⁸⁶ Para Balkin os valores-chave da liberdade de expressão: “[...] a proteção da liberdade individual de expressar ideias, formar opiniões, criar arte e se envolver em pesquisas; a capacidade de indivíduos e grupos de compartilhar seus pontos de vista com os outros e construir sobre as idéias de outros; e a promoção e disseminação de conhecimento e opinião”. Tradução livre do original: “[...] the protection of individual freedom to express ideas, form opinions, create art, and engage in research; the ability of individuals and groups to share their views with others, and build on the ideas of others; and the promotion and dissemination of knowledge and opinion.” (BALKIN, Jack M. The Future of Free Expression in a Digital Age, **Pepperdine Law Review**, v. 36, Iss. 2, 2009. p. 427. Disponível em: <https://digitalcommons.pepperdine.edu/plr/vol36/iss2/9/>. Acesso em: 04/11/2021).

¹⁸⁷ BALKIN, Jack M. The Future of Free Expression in a Digital Age, **Pepperdine Law Review**, v. 36, Iss. 2, 2009. p. 427. Disponível em: <https://digitalcommons.pepperdine.edu/plr/vol36/iss2/9/>. Acesso em: 04/11/2021.

a maneira como devemos pensar sobre a Primeira Emenda e os princípios da liberdade de expressão.”¹⁸⁸ e passando a entender o objetivo do princípio da liberdade de expressão como promotor, não apenas da democracia, mas também, de forma mais ampla, denominada de cultura democrática por Balkin e incluindo, nesse contexto a participação no autogoverno representativo. O conteúdo das redes digitais oferece os produtos do que as pessoas fazem na *Internet* e transcende o estado-nação; os novos formadores de opinião participam de discussões, debates e atividades coletivas, os quais transcendem as fronteiras nacionais.

Nessa inflexão, percebe-se que a liberdade de expressão de cunho individual encontra resistência perante o mercado, ou seja, nas plataformas digitais, posto que somente proporcionar um palco de exposição de ideias, conteúdo, conhecimento, informação não gera o consumo de bens de mídia. Nesse ponto, as empresas de redes sociais, tendem a não defender a liberdade de expressão de forma plena, tendendo a direcionar sua utilização para fortalecer consumidores.

As empresas de tecnologia digital desenvolvem meios de restringir e canalizar a participação democrática, com discriminação de pacotes e gerenciamento de direitos digitais, sob a justificativa interpretativa da liberdade de expressão que vincula a expressão aos direitos de propriedade.¹⁸⁹

Essa preocupação representa uma tensão na teoria da liberdade de expressão do século XXI, pois, como apontado, as novas oferecem aos cidadãos uma gama de oportunidades, contudo, essas mesmas tecnologias também tornam a informação e a cultura em mercadorias cada vez mais valiosas, podendo ser compradas, vendidas e exportadas para os mercados do mundo inteiro.

Uma das funções centrais das mídias sociais, observada por Balkin, é a realização da curadoria da opinião pública, não apenas por meio de *feeds* e resultados de pesquisa individualizados, mas também por meio da aplicação de *padrões e termos de serviço da*

¹⁸⁸ Tradução livre do original: “The rise of digital networks as a dominant technology for speech in our age transforms the way we should think about the First Amendment and the principles of freedom of expression”. (BALKIN, Jack M. The Future of Free Expression in a Digital Age, **Pepperdine Law Review**, v. 36, Iss. 2, 2009. p. 438. Disponível em: <https://digitalcommons.pepperdine.edu/plr/vol36/iss2/9/>. Acesso em: 04/11/2021).

¹⁸⁹ BALKIN anota que essa “concepção capitalista é importante tanto para explicar o que é liberdade de expressão (liberdade de regulamentação de negócios) e o que não é (um limite aplicável à expansão dos direitos de propriedade intelectual) [...]” Tradução livre do original: “The developing capitalist conception of freedom of speech (and its accompanying denial of free speech limitations on the growth of intellectual property) [...]” (BALKIN, Jack M., Digital Speech and Democratic Culture: A Theory of Freedom of Expression for the Information Society, **New York University Law Review**, v. 79, p. 1–55, 2004. p. 46. Disponível em: https://openyls.law.yale.edu/bitstream/handle/20.500.13051/1712/Digital_Speech_and_Democratic_Culture_A_Theory_of_Freedom_of_Expression.pdf?sequence=2. Acesso em: 04/11/2021).

comunidade. As redes sociais fazem a curadoria não apenas retirando ou reorganizando o conteúdo, mas também regulando a velocidade de propagação e o alcance do conteúdo.¹⁹⁰

De modo geral, o princípio da liberdade de expressão permite ao Estado impor apenas um conjunto muito limitado de normas de civilidade ao discurso público, deixando as instituições intermediárias livres para instituir normas mais rígidas de acordo com seus valores.¹⁹¹

¹⁹⁰ BALKIN, Jack M., How to Regulate (and Not Regulate) Social Media. **Journal of Free Speech Law**. v. 1, p.71-96, Knight Institute Occasional Paper Series, No. 1, Yale Law School, Public Law Research Paper. 2021. Disponível em: <https://ssrn.com/abstract=3484114>. Acesso em: 10/12/2021. p. 75. Disponível em: https://openyls.law.yale.edu/bitstream/handle/20.500.13051/1712/Digital_Speech_and_Democratic_Culture_A_Theory_of_Freedom_of_Expression.pdf?sequence=2. Acesso em: 04/11/2021.

¹⁹¹ BALKIN, Jack M., How to Regulate (and Not Regulate) Social Media. **Journal of Free Speech Law**. v. 1, p.71-96, Knight Institute Occasional Paper Series, No. 1, Yale Law School, Public Law Research Paper. 2021. Disponível em: <https://ssrn.com/abstract=3484114>. Acesso em: 10/12/2021. p. 76. Disponível em: https://openyls.law.yale.edu/bitstream/handle/20.500.13051/1712/Digital_Speech_and_Democratic_Culture_A_Theory_of_Freedom_of_Expression.pdf?sequence=2. Acesso em: 04/11/2021.

CAPÍTULO II – GOVERNANÇA PRIVADA E LIBERDADE DE EXPRESSÃO NAS PLATAFORMAS DE MÍDIA SOCIAL

Como visto, com a popularização da *Internet* na década de 90, passou a fazer parte da vida hodierna do cidadão, que agora está conectado e com imensa facilidade para formação de opinião, aprendizagem, comunicação, socialização, acesso a diversos serviços públicos e privados, além de permitir desenvolver a criatividade com a produção de conteúdo *online*.¹⁹²

Nessa realidade, ofertou-se ao indivíduo cada vez mais oportunidades dependentes, diretamente, do acesso à *Internet*. Por outro lado, exigiu-se esforço maior de compreensão do usuário para evitar os riscos, devendo o uso a ser seguro e proveitoso. Destarte, termos como governança e regulação se tornaram tão poderosos e essenciais em nossas vidas, deixando de ser assunto exclusivo dos profissionais e autoridades, tornando-se tema de debate hodierno entre o público em geral. Tal popularização evidenciou-se devido à atuação de atores privados de plataformas de redes sociais, dos posicionamentos de governos, bem como das proposições legislativas.

A governança privada representa um dos lados do triângulo, na concepção de Balkin, a liberdade de expressão é um triângulo.¹⁹³ Ocupante dessa estrutura, Balkin aponta que “as empresas de mídia social criam sistemas complexos de governança privada e burocracia

¹⁹² Por outro lado, reconhecemos o problema da exclusão digital, da desigualdade digital, além do analfabetismo digital. Nesse ponto, ver o importante estudo de Sofia Ranchordás, que argumenta: “À medida que surgem novas brechas digitais, a desigualdade digital não apenas reproduz desigualdades socioeconômicas de longa data, mas também se coloca como uma fonte autônoma de exclusão” e conclui: “Enquanto muitos cidadãos simplesmente não estão conectados à Internet, outros não têm acesso significativo a ela, não têm a alfabetização (digital) para poder abraçar plenamente seu potencial e se envolver com o crescente número de serviços públicos digitais ou não estão dispostos para adotar as tecnologias digitais”. Tradução livre do original: “As new digital divides emerge, digital inequality is not only reproducing longstanding socioeconomic inequalities but it is also placing itself as a standalone source of exclusion.” [...] “While many citizens simply are not connected to the Internet, others do not have meaningful access to it, do not have the (digital) literacy to be fully able to embrace its potential and engage with the growing number of digital public services or are not willing to uptake digital technologies.” (RANCHORDAS, Sofia, *The Digitalization of Government and Digital Exclusion: Setting the Scene* (29 de abril de 2020), a ser publicado em FERREIRA MENDES, G. & BLANCO DE MORAIS, C. (Org.). *Direito Público e Internet: Democracia, Redes Sociais e Regulação do Ciberespaço*, FGV Publicações/IDP/Univ. Lisboa, (Public Governance 4.0, 2020). University of Groningen, **Faculty of Law Research Paper**, N. 30/2020. Disponível em: <https://ssrn.com/abstract=3663051> ou <http://dx.doi.org/10.2139/ssrn.3663051>. Acesso em: 28/10/2021). Para uma visão do tema na América Latina ver: BELLO, Pablo; SASTRE, Andrés. Repensar as políticas públicas para fechar o fosso digital na América Latina. In. BELLI, Luca; CAVALLI, Olga (Org.). **Governança e regulações da Internet na América Latina: análise sobre infraestrutura, privacidade, cibersegurança e evoluções tecnológicas em homenagem aos dez anos da South School on Internet Governance**. Rio de Janeiro: Escola de Direito do Rio de Janeiro da Fundação Getúlio Vargas, 2019. p. 247-266.

¹⁹³ BALKIN, Jack M. Free Speech is a triangle. **Columbia Law Review**. v. 118, p. 2011–2055, 2018. Disponível em: <https://openyls.law.yale.edu/bitstream/handle/20.500.13051/17953/Balkin%2C%20Free%20Speech%20is%20a%20Triangle.pdf?sequence=1&isAllowed=y>. Acesso em: 04/11/2021.

privada que governam os usuários finais de forma arbitrária e sem o devido processo e transparência.”¹⁹⁴

Destarte, esse capítulo objetiva analisar como as empresas de mídia social devem equilibrar a observância dos direitos fundamentais em suas plataformas, particularmente, em relação à liberdade de expressão. Primeiramente, apresentando a definição de governança da *Internet*; ao analisar os principais atores numa dimensão subjetiva, com ênfase na visão de governança privada, realizada por empresas de mídia social; na dimensão objetiva, examinando, em linhas gerais, a moderação de conteúdo como uma das práticas de governança privada; após, passamos a abordar a governança de plataformas e direitos humanos, analisando o que a doutrina vem denominando de “tom constitucional dos termos de uso”; no final, examinamos o efeito horizontal da proteção da liberdade de expressão.

2.1. Governança (privada) da *internet*: definição.

A governança privada, representa uma camada¹⁹⁵ do conceito mais amplo da denominada governança da *Internet*, numa ordem do conceito específico para o mais amplo. Ademais, afasta-se eventual confusão quanto à semelhança entre os verbetes governo e governança. Governo possui uma estrutura centralizada, com estratégias as quais assumem uma forma concreta, composto por uma autoridade única, com poder forte e coercitivo. Enquanto a governança possui natureza de diálogo, de comum acordo.¹⁹⁶

Nessa perspectiva, Gonçalves e Costa entendem governança realizada por meio de diálogos firmados entre partes diversas apresentando um consenso, sem, contudo, exigir a necessidade de uma autoridade central ou obrigação legal.¹⁹⁷ Outra característica da estrutura

¹⁹⁴ Tradução livre do original: “Second, social media companies create complex systems of private governance and private bureaucracy that govern end users arbitrarily and without due process and transparency” (BALKIN, Jack M. Free Speech is a triangle. **Columbia Law Review**. v. 118, p. 2011–2055, 2018. p. 2011. Disponível em: <https://openyls.law.yale.edu/bitstream/handle/20.500.13051/17953/Balkin%2C%20Free%20Speech%20is%20a%20Triangle.pdf?sequence=1&isAllowed=y>. Acesso em: 04/11/2021).

¹⁹⁵ “Na camada superior – das aplicações de Internet – vigora uma espécie de governança privada, cujas regras são definidas, em grande medida, de forma específica pelos provedores dos diferentes serviços disponibilizados aos usuários finais, como se identifica claramente nos termos e condições de uso de redes sociais de escala global.” (WAGNER, Flávio Rech; CANABARRO, Diego Rafael. A governança da *Internet*: definição, desafios e perspectivas. In: PIMENTA, Marcelo Soares; CANABARRO, Diego Rafael (Org.). **Governança Digital**. Porto Alegre: Editora da UFRGS/CEGOV, 2014. p. 199).

¹⁹⁶ HUIDOBRA, Marina Stephanie Ramos; OLIVEIRA, Natalia Rosa de. Governança global: o comitê das nações unidas para o uso pacífico do espaço (copuos) como mecanismo de governança global espacial. In: BENEVIDES, Solon Henriques de Sá e; GAUDÊNCIO, Francisco de Sales (Org.). **Governança e Direitos Fundamentais: revisitando o debate entre o público e o privado**. Porto: Instituto Iberoamericano de Estudos Jurídicos, 2020. p. 201.

¹⁹⁷ GONÇALVES, Alcindo; COSTA, José A. Fontoura. **Governança Global e Regimes Internacionais**. 1 ed. São Paulo: Editora Almedina, 2011. p. 50.

de governança é a geração de efeitos por meio de instrumentos geralmente tangíveis, por vezes normatizados, que perfilham a complexa fronteira entre o público e o privado.

O termo “governança” possui aplicação ampla em diversas áreas do conhecimento, sendo que a maior parte de suas aplicações não possui ligação com a ideia centralizada e hierárquica estatal. Com isso, a governança tem amplitude para incidir em campos públicos, privados ou multissetoriais.¹⁹⁸

No campo da ciência política, encontra-se o uso adequado do termo governança empregado nas relações de gestão configurando a *Internet*. De acordo com Guy Peters, a governança é como um processo de coordenação de múltiplas ações coletivas as quais circundam tanto governos, como agentes privados e transnacionais repercutindo na vida política e econômica, assumindo significativa importância perante os desafios cujas ações individuais não são capazes de suportar, por exemplo, escassez de recursos, controle do tráfico de armas, mudança climática. Contudo, Peters aponta para essa noção indissociável do conceito de Estado, tendo em vista que governança implica, de igual modo, a adoção de alguma conceição de *accountability*.¹⁹⁹

Nesse sentido, David Levi-Faur ao discutir a governança como uma agenda de pesquisa interdisciplinar sobre ordem e desordem, eficiência e legitimidade, no contexto da hibridização de modos de controle permitindo a produção de uma ordem fragmentada e multidimensional dentro do Estado, pelo Estado, sem o Estado e além do Estado, pontua a importância de tratar a governança como um sistema cada vez mais híbrido, compreendendo diversos tipos de composição regulatórias.²⁰⁰

¹⁹⁸ Keller aponta a diversificação de sua aplicação, no direito societário, como governança corporativa consistindo no sistema em que as sociedades empresárias são administradas e controladas. Na perspectiva econômica, a governança assume papel no espectro das ciências e estudos da boa ordem e dos arranjos institucionais funcionais. (KELLER, Clara Iglesias. **Regulação nacional de serviços na internet**: exceção, legitimidade e o papel do Estado. Rio de Janeiro: Lumen Juris, 2019. p.78-79).

¹⁹⁹ PETERS, Guy. Governance As Political Theory. **Oxford Handbooks On line**. Oxford University Press. 29 de março de 2012, p. 02-03, Disponível em: <https://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199560530.001.0001/oxfordhb-9780199560530-e-2>. Acesso em: 10/05/2022.

²⁰⁰ O Autor completa que: “Em suma, estamos no apogeu da “Big Governance” e a grande questão da governança, ou seja, a estratégia de projetos e controle de governança, é determinar não qual modo puro de governança é mais eficaz ou mais legítimo, mas quais híbridos são. Precisamos conceituar uma ordem mundial onde a governança seja cada vez mais um híbrido de diferentes sistemas de controle regulatório; onde a regulação estatal co-evolui com a regulação civil; a regulamentação nacional se expande com a regulamentação internacional e global; a regulação privada co-evolui e se expande com a regulação pública; a regulação empresarial co-evolui com a regulação social; as regulamentações voluntárias se expandem com as coercitivas; e o próprio mercado é utilizado ou mobilizado como mecanismo regulatório.” Tradução livre do original: “In short, we are in the heyday of “Big Governance” and the major question of governancing, that is, the strategy of governance designs and control, is to determine not which pure mode of governance is more effective or more legitimate but which hybrids are. We need to conceptualize a world order where governance is increasingly a hybrid of different systems of regulatory control; where statist regulation co-evolves with civil regulation;

Como afirmado, no aspecto semântico da ciência política compreende-se a complexidade das nuances que compõem a governança da *Internet*, como suas estruturas, práticas e agentes. Conforme preceitua Keller, constitui um exercício de funções, apontando cinco aspectos definidores do conceito: abordagem pluricêntrica (1), ausência de monocracia e hierarquia (2); privilegia processos e funções de governos, e não estruturas de governo (3); abordagens são normativas, preconizando um ideal e uma realidade empírica.²⁰¹ Com espeque nesses parâmetros, tem-se que se trata de um arranjo policêntrico, onde não há designação de um gestor principal que reúna todas as funções regulatórias. É dizer, não há hierarquização entre os atores, muito menos uma estrutura de governo formalmente constituída e centralizada.²⁰²

Noutro passo, conforme previsão do parágrafo 34 da Agenda de Túnis para a Sociedade da Informação, acolhido durante a Cúpula Mundial sobre a Sociedade da Informação da ONU, a *World Summit on the Information Society* (WSIS), forneceu a seguinte definição:²⁰³

A governança da Internet é o desenvolvimento e aplicação pelos governos, setor privado e sociedade civil, no desempenho de suas respectivas funções, princípios, normas, regras, procedimentos de tomada de decisão e programas comuns que moldam a evolução e uso da Internet”.²⁰⁴

national regulation expands with international and global regulation; private regulation co- evolves and expands with public regulation; business regulation co-evolves with social regulation; voluntary regulations expand with coercive ones; and the market itself is used or mobilized as a regulatory mechanism.” (LEVI-FAUR, David. From “Big Government” to “Big Governance”? *Jerusalem Papers in Regulation & Governance*. n. 35, julho de 2011, p. 17. Disponível em: <http://regulation.huji.ac.il/papers/jp35.pdf>. Acesso em: 10/05/2022)

²⁰¹ KELLER, Clara Iglesias. **Regulação nacional de serviços na internet: exceção, legitimidade e o papel do Estado**. Rio de Janeiro: Lumen Juris, 2019. p. 81.

²⁰² Nessa estrutura, pontua Clara Keller, que: “Empresas privadas de diferentes tamanhos e atividades, governos, organizações paraestatais de governança multisetorial e até provedores de acesso e prestadores de serviços de diferentes portes influenciam esse sistema – seja através de participação em fórum com espoco de instruir ou definir política pública, seja através do exercício natural de suas atividades (código).” (KELLER, Clara Iglesias. **Regulação nacional de serviços na internet: exceção, legitimidade e o papel do Estado**. Rio de Janeiro: Lumen Juris, 2019. p.82).

²⁰³ WSIS é a abreviação de Cúpula Mundial sobre a Sociedade da Informação. O Summit é popularmente conhecido como WSIS (pronuncia-se wiss-iss). Previsto desde seu início em duas fases, a primeira fase da Cúpula foi realizada em Genebra, de 10 a 12 de dezembro de 2003, onde as bases foram lançadas por meio de um acordo sobre uma Declaração de Princípios e um Plano de Ação. A segunda fase foi realizada em Túnis, de 16 a 18 de novembro de 2005, com a adoção do Compromisso de Túnis e da Agenda de Túnis para a Sociedade da Informação. Este último inclui capítulos sobre mecanismos de financiamento, governança da *Internet* e implementação e acompanhamento. Disponível em: https://www.itu.int/net/wsis/basic/faqsanswer.asp?lang=en&faq_id=46. Acesso em: 20/05/2022.

²⁰⁴ Tradução livre do original: “A working definition of Internet governance is the development and application by governments, the private sector and civil society, in their respective roles, of shared principles, norms, rules, decision-making procedures, and programmes that shape the evolution and use of the Internet.” (World Summit on the Information Society. **Tunis agenda for the information Society**. 2005. Disponível em: <https://www.itu.int/net/wsis/docs2/tunis/off/6rev1.html>. Acesso: 15/01/2022.

Tal definição “oficial”, apesar de apresentar um norte e demonstrar a complexidade envolvida na governança da *internet*.²⁰⁵ Parte dos estudiosos do tema, entendem esse conceito como carente de trabalhos acadêmicos mais aprofundados,²⁰⁶ não sendo capaz de apresentar soluções e paradigmas para o futuro da *internet* e seus diversos desafios. Francesca Musiani, argumentando que “essa definição talvez seja muito ampla para ser útil no nível ‘operacional’, delimitando precisamente as fronteiras do campo, seu caráter inclusivo, no entanto, torna muito consensual e fornece vários elementos que, embora discutíveis, são de alguma utilidade.”²⁰⁷

Pimentel e Canabarro, pontuam que o conceito “deixa em aberto o rol de questões que giram em torno da evolução e do uso da *Internet* na atualidade.”²⁰⁸, sendo assim, apenas a apresentação de um recorte da situação vivenciada na época dos debates na ONU, merecendo mais aprofundamento e aperfeiçoamento para enfrentar os problemas atuais no ecossistema digital.

Como lembram Hofmann, Katnbach e Gollatz, a WSIS conseguiu atrair novos especialistas estudiosos sobre o tema, servindo para fomentar a busca pela primeira definição de governança da *internet*. Contudo, tal celeridade resultou numa disputa política, e não acadêmica. Como aduzem os autores:

A definição do WGIG é uma criatura de seu tempo; é um artefato de múltiplos atores dissociado da pesquisa acadêmica sobre a natureza dos arranjos de governança transnacional. No entanto, com base no trabalho realizado pelo WGIG, a

²⁰⁵ Luca Belli, aponta como mérito dessa definição, evitar conflitos epistemológicos sobre o que pode ser definido como governança: “Desde a década de 1980, o conceito de governança generalizou-se e a versatilidade de seu uso sugeriu a existência de quase tantos conceitos de governança quanto pesquisadores no campo. Em contraste com o termo “governo”, que caracteriza a instituição governante, o termo “governança” é geralmente usado para se referir ao aspecto processual e, particularmente, ao modo de governar um sistema. Esse conceito tem sido usado desde a Idade Média para descrever o sistema político caracterizado pela participação institucionalizada de diferentes grupos corporativos, como a Igreja, os atores econômicos e os poderes territoriais. Desde a Idade Média, muitas versões do conceito de governança foram elaboradas e promovidas.” (BELLI, Luca. Governança e regulações da *internet*: uma apresentação crítica. In: BELLI, Luca; CAVALLI, Olga (Org.). **Governança e regulações da Internet na América Latina**: análise sobre infraestrutura, privacidade, cibersegurança e evoluções tecnológicas em homenagem aos dez anos da South School on Internet Governance. Rio de Janeiro: Escola de Direito do Rio de Janeiro da Fundação Getúlio Vargas, 2019. p. 45).

²⁰⁶ Ver a propósito: HOFMANN, Jeanette; KATZENBACH, Christian; GOLLATZ Kirsten. Between coordination and regulation: Finding the governance in Internet governance. **New Media & Society**. v. 19, n. 9, p. 406-1423, 2017. Disponível em: doi:10.1177/1461444816639975.

²⁰⁷ Tradução livre do original: “[...] cette définition est peut-être trop large pour être utile sur le plan «opérationnel» en délimitant précisément les frontières du domaine, son caractère inclusif la rend néanmoins très consensuelle et elle fournit plusieurs éléments qui, bien qu’ils soient discutables, sont d’une certaine utilité.” (MUSIANI, Francesca. La gouvernance d’Internet, brève histoire d’un espace à controverses, **Terminal**, p. 132-133, 2022, Disponível: <http://journals.openedition.org/terminal/8398>; ou DOI: <https://doi.org/10.4000/terminal.8398>. Acesso em: 25/05/2022).

²⁰⁸ WAGNER, Flávio Rech; CANABARRO, Diego Rafael. A governança da *Internet*: definição, desafios e perspectivas. In: PIMENTA, Marcelo Soares; CANABARRO, Diego Rafael (Org.). **Governança Digital**. Porto Alegre: Editora da UFRGS/CEGOV, 2014. p. 193.

última década viu um aumento na pesquisa sobre as características da governança da Internet.²⁰⁹

Numa visão mais simples e direta, Mueller, define governança da *internet* como conjunto dinâmico de disputas e deliberações sobre como a *Internet* é coordenada, administrada e moldada para refletir políticas públicas.²¹⁰

Nesse mesmo raciocínio, Mats, Oermann e Schlz, tratando sobre o desenvolvimento e aplicação de princípios, normas, regras, procedimentos de tomada de decisão e programas compartilhados que moldam a evolução e o uso da *internet*. Aduzem a governança da *Internet* “parece ser uma cifra adequada para descrever os processos abertos de normatização e as estruturas normativas evoluindo neste domínio que reúnem vários tipos de atores.”²¹¹ Para os autores, essa cifra é utilizada para identificar estruturas e processos em níveis diferentes de comunicação na *internet*, podendo ser distinguidos analiticamente:

no nível de conteúdos e serviços informacionais, que os usuários nas extremidades das redes interligadas da internet implementam entre si; ao nível dos protocolos e outros requisitos lógicos, habilitando os serviços e aplicações ao nível do conteúdo; e ao nível da infraestrutura física, os cabos e frequências através dos quais os dados necessários são transferidos.²¹²

Acrescente-se, ainda, que o estudo realizado pela WSIS, revelou-se carente de esclarecimento de como se desenvolveria, e, como se implementaria concretamente a cooperação entre diferentes *stakeholders*.²¹³ Deixando um vácuo para especulação sobre a

²⁰⁹ HOFMANN, Jeanette; KATZENBACH, Christian; GOLLATZ Kirsten. Between coordination and regulation: Finding the governance in Internet governance. **New Media & Society**. v. 19, n. 9, p. 1406-1423, 2017. p. 1408. Disponível em: doi:10.1177/1461444816639975. Acesso em 20/05/2022.

²¹⁰ MUELLER, Milton L. A Battle for the Soul of the Internet. **Networks and States: The Global Politics of Internet Governance**, MIT Press, 2010, p. 1-14. p. 9. Disponível em: www.jstor.org/stable/j.ctt5hhcz0.3. Acesso em 29/05/2022.

²¹¹ MAST, Tobias; OERMANN, Markus; SCHULZ, Wolfgang. Doing Internet Governance: Constructing Normative Structures Inside and Outside of Intermediary Organisations. **GigaNet: Global Internet Governance Academic Network**, Annual Symposium 2016. p. 3. Disponível em: https://ssrn.com/abstract=2909367 ou http://dx.doi.org/10.2139/ssrn.2909367. Acesso em 01/06/2022.

²¹² Tradução livre do original: “[...] at the level of informational contents and services, which the users at the ends of the cross-linked networks of the internet implement among themselves; at the level of protocols and other logical requirements, enabling the services and applications on content level; and at the level of the physical infrastructure, the cables and frequencies via which the necessary data transfers”. (MAST, Tobias; OERMANN, Markus; SCHULZ, Wolfgang. Doing Internet Governance: Constructing Normative Structures Inside and Outside of Intermediary Organisations. **GigaNet: Global Internet Governance Academic Network**, Annual Symposium 2016. p. 3. Disponível em: https://ssrn.com/abstract=2909367 ou http://dx.doi.org/10.2139/ssrn.2909367. Acesso em 01/06/2022).

²¹³ “O termo inglês stakeholder define qualquer pessoa física ou jurídica envolvida nos processos deliberativos e de tomada de decisão que ocorrem nos níveis internacional, nacional ou local. A participação das partes interessadas é justificada por seu interesse ou stake em um determinado processo.” (BELLI, Luca. Governança e regulações da internet: uma apresentação crítica. In: BELLI, Luca; CAVALLI, Olga (Org.). **Governança e regulações da Internet na América Latina: análise sobre infraestrutura, privacidade, cibersegurança e evoluções tecnológicas em homenagem aos dez anos da South School on Internet Governance**. Rio de Janeiro: Escola de Direito do Rio de Janeiro da Fundação Getúlio Vargas, 2019. p. 44). Conferir também, Wagner e Canabarro, que afirmam: “No mesmo documento em que definiu a governança da Internet, a ONU reconheceu

caracterização como um processo *multistakeholder* e, o mais importante, poderia ser apontando como uma forma aberta, inclusiva ou democrática de *multistakeholderismo*.²¹⁴

Com isso, desde então diversos estudos na área da governança da *Internet* passaram a se dedicar sobre os processos e instituições classificadas como *multistakeholder*, sobretudo quanto ao modelo *multistakeholder* criado na área do sistema de nomes de domínio – o *Domain Name System* – DNS, e dos protocolos de *internet* – TCP/IP, ambos componentes da estrutura técnica do funcionamento da *internet*.

Considerando a densidade de formas e contextos de abordagem da governança da *internet* – a qual, em síntese, foi descrita como sendo tanto à governança da Rede em si, quanto de todas as coisas as quais a rodeiam e com ela se relacionam²¹⁵ –, como meio de facilitar didaticamente a compreensão, bem como direcionar o estudo proposto nesta dissertação, adotamos a abordagem da governança em dois sentidos: subjetivo e objetivo, proposta por Clara Iglesias Keller.²¹⁶

2.2. Governança (privada) da *internet*: dimensão subjetiva

Nessa visão, a dimensão subjetiva da governança da *Internet* compõe o *grupo de atores* que integram os processos e gestão que influenciam na formação de políticas públicas no ambiente *online*. Dentre eles, incluem governos nacionais, entidades supranacionais, organizações da sociedade civil, empresas privadas e fóruns privados multissetoriais.

Nesse ponto, tais atores representam instituições criadas para gerenciar a chamada “raiz” da *internet*,²¹⁷ o qual é o conjunto centralizado de servidores, onde é armazenado a base de dados com todos os endereços alfanuméricos, identificando todas suas redes integrantes e,

como stakeholders na governança da Internet os governos, o setor privado, a sociedade civil (entendida como o terceiro setor), organizações internacionais e a academia. Em síntese apertada, esse grupo bastante difuso e abrangente compõe a ‘comunidade da Internet’. Deve-se notar, porém, que cada um desses grupos de atores varia consideravelmente em sua natureza. [...]” (WAGNER, Flávio Rech; CANABARRO, Diego Rafael. A governança da Internet: definição, desafios e perspectivas. In: PIMENTA, Marcelo Soares; CANABARRO, Diego Rafael (Org.). **Governança Digital**. Porto Alegre: Editora da UFRGS/CEGOV, 2014. p. 196)

²¹⁴ BELLI, Luca. Governança e regulações da internet: uma apresentação crítica. In: BELLI, Luca; CAVALLI, Olga (Org.). **Governança e regulações da Internet na América Latina**: análise sobre infraestrutura, privacidade, cibersegurança e evoluções tecnológicas em homenagem aos dez anos da South School on Internet Governance. Rio de Janeiro: Escola de Direito do Rio de Janeiro da Fundação Getúlio Vargas, 2019. p. 45.

²¹⁵ CANABARRO, Diego. Governança Global da Internet: Aspectos Conceituais, Questões da Agenda Contemporânea e Prospectos para o Estudo do Tema. In: OPPERMANN, Daniel. (Ed.). **Internet Governance in the Global South: History, Theory, and Contemporary Debates**. São Paulo: University of São Paulo, 2018. p. 76. Disponível em: <https://nbn-resolving.org/urn:nbn:de:0168-ssor-65805-2>. Acesso em: 08/07/2022.

²¹⁶ KELLER, Clara Iglesias. **Regulação nacional de serviços na internet**: exceção, legitimidade e o papel do Estado. Rio de Janeiro: Lumen Juris, 2019. p. 85.

²¹⁷ Conferir: MUELLER, Milton. **Ruling the Root**: Internet Governance and the Taming of Cyberspace. Cambridge, MA: MIT Press, 2002.

com isso, é capaz de identificar cada terminal conectado em suas pontas. Essa identificação é realizada pelo chamado *Internet Protocol* (IP), responsável pelo endereço numérico, e no *Domain Name System* (DNS), que confere o correspondente alfabético destinado a facilitar a memorização de endereços na rede pelos usuários humanos.²¹⁸

Resumidamente, abordaremos as principais entidades e organizações, que possuem e/ou possuíram participação relevante no desenvolvimento da governança, dentre elas: *Institute of Electrical and Electronics Engineers* (IEEE), *International Telecommunication Union* (ITU), *Internet Engineering Task Force* (IETF), *World Wide Web Consortium* (W3C), *Internet Assigned Number Authority* (IANA), *Internet Corporation for Assigned Names and Numbers* (ICANN), *World Summit on the Information Society* (WSIS), *Working Group on Internet Governance* (WGIG) e *Internet Governance Forum* (IGF).

- *Institute of Electrical and Electronics Engineers* (IEEE) e *International Telecommunication Union* (ITU). Trata-se de organizações engajadas nos padrões utilizados nas telecomunicações, possuindo atuação em parte da infraestrutura da *Internet*. A União Internacional de Telecomunicações (ITU), foi criada em 1896, sendo responsável pela padronização na área de telecomunicações. Nessa mesma linha, o IEEE tem a incumbência, dentre outras, pelos padrões elétricos das interfaces de comunicação. Saliente-se, que ambas não criavam padrões exclusivamente para a *internet*, inclusive, precedendo a própria *Internet*.

Em 1947, a ITU foi diretamente ligada à ONU²¹⁹, com a natureza de Agência Especializada, responsável pela gestão global das telecomunicações. Mas, recentemente, tem-se aproximado em torno das Redes de Nova Geração, as *New Generations Networks* (NGNs)²²⁰, representa uma evolução para rede mundial. Esta aproximação da ITU com a *Internet* é objeto de críticas, sob o pretexto de a configuração poder-se dar, tal como é a relação entre governos e as empresas de telecomunicações, desprezando o caráter colaborativo e aberto dos padrões da *Internet*, inclusive, a participação dos setores

²¹⁸ WAGNER, Flávio Rech; CANABARRO, Diego Rafael. A governança da Internet: definição, desafios e perspectivas. In: PIMENTA, Marcelo Soares; CANABARRO, Diego Rafael (Org.). **Governança Digital**. Porto Alegre: Editora da UFRGS/CEGOV, 2014. p. 196.

²¹⁹ “Como é uma organização ligada à ONU, países podem ser seus membros. Hoje a ITU tem cerca de 191 governos como membros. Além disso, organizações do setor privado também podem ser membros da organização. Hoje há mais de 700 empresas nessa categoria.” GATTO, Raquel F.; MOREIRAS, Antonio M.; e GETSCHKO, Demi. Governança da Internet: conceitos, evolução e abrangência. In. **27º Simpósio Brasileiro de Redes de Computadores e sistemas Distribuídos**, 2009. p. 77.

²²⁰ “As forças tecnológicas e de mercado estão levando operadoras de rede e provedores de serviços de comunicação eletrônica em todo o mundo a migrar suas redes para uma base de Protocolo de Internet (IP). Na maior parte do mundo, a rede, as redes resultantes são referidas como redes de próxima geração (NGNs).” (MARCUS, J. Scott; ELIXMANN, Dieter. **Regulatory Approaches to Next Generation Networks (NGNs): An International Comparison**. TPRC, 2007. Disponível em: <https://ssrn.com/abstract=2118249>. Acesso em 05/06/2022).

interessados, na governança²²¹. Anote-se que a ITU compôs as organizações as quais propuseram a Cúpula Mundial sobre a Sociedade da Informação, inclusive, ainda, verifica o cumprimento dos deveres esculpidos em Túnis (2005).

O *Institute of Electrical and Electronics Engineers* (IEEE) foi criado em 1884, nos Estados Unidos, constituído como uma sociedade técnico-profissional internacional, sem fins lucrativos, centrado no avanço teórico-prático da engenharia elétrica, eletrônica e da computação. Possui aproximadamente 375 mil os associados, distribuído em cerca de 160 países. Um dos principais padrões produzidos é a IEEE 802.3 (redes *ethernet*), a IEEE 802.11 (*wi-fi*) e a IEEE 802.15.4 (*wireless personal area networks*). As normas criadas pelo Instituto são restritas aos associados, contudo, é possível o acesso por não-associados desde que paguem pelo serviço.

- *Internet Engineering Task Force* (IETF): originou a partir de um fluxo de debates escritos, realizado por um grupo de pesquisadores denominado *Network Working Group* (NWG), os quais tratavam, primordialmente, sobre o desenvolvimento, além de análises de propostas para as redes, após algumas etapas transitórias, o NWG originou a IETF.²²² Esses processos de deliberação e resolução de demandas de interconectividade de sistemas autônomos, são condensados numa série de documentos denominados *Request for Comments* (RFC) ou Requisição de Comentários. Chenou e Fuerte sintetizam que “apesar do nome, compreende documentos técnicos e de padronização da *Internet*. As RFCs às vezes são chamadas de “leis” da *Internet*.”²²³

Na prática, quando surge uma determinada questão ou proposta ela circula entre a chamada “comunidade da *Internet*” até que, após debates, deliberações, proposições, reformas, esteja madura para então ter condições de ser adotada como um padrão consolidado e, assim, torne-se majoritariamente aceito nas rotinas de operação da rede.²²⁴

A IETF é custeada e mantida pela *The Internet Society* (ISOC), instituição criada como uma organização internacional, sem fins lucrativos, tem como objetivo colaborar com a

²²¹ GATTO, Raquel F.; MOREIRAS, Antonio M.; e GETSCHKO, Demi. Governança da Internet: conceitos, evolução e abrangência. In: 27º **Simpósio Brasileiro de Redes de Computadores e sistemas Distribuídos**, 2009. p. 77.

²²² GATTO, Raquel F.; MOREIRAS, Antonio M.; e GETSCHKO, Demi. Governança da Internet: conceitos, evolução e abrangência. In: 27º **Simpósio Brasileiro de Redes de Computadores e sistemas Distribuídos**, 2009. p. 79.

²²³ CHENOU, Jean-Marie; FUERTE, Juan Sebastián Rojas. The Difficult Path to the Insertion of the Global South in Internet Governance. In: OPPERMAN, Daniel. (Ed.). **Internet Governance in the Global South: History, Theory, and Contemporary Debates**. São Paulo: University of São Paulo, 2018. p. 57. Disponível em: <https://nbn-resolving.org/urn:nbn:de:0168-ssor-65805-2>. Acesso em: 08/07/2022.

²²⁴ Todos *Request for Comments* encontram-se disponíveis no site: <https://www.rfc-editor.org/rfc-index.html>. Acesso em 11/06/2022.

primazia do desenvolvimento de padrões de *Internet*, além de aspectos educacionais e de regulação da rede mundial. Possui, ainda, como princípio a manutenção de rede aberta, com acesso universal.

- *World Wide Web Consortium (W3C)*, foi criada em 1994, por Tim Berners-Lee, tendo como objetivo garantir a compatibilidade entre várias tecnologias surgidas e são incorporadas à *Web*. Saliente-se que a W3C tem característica de aplicação, não sendo essencial para o funcionamento em si da *Internet*, contudo, é importante na interoperabilidade da *Web*. Ademais, possui natureza de consórcio integrado por organizações de múltiplas naturezas, como representantes do setor público, além de outros membros.

A W3C possui cerca de 110 padrões publicados, chamado de “recomendações W3C”, diferente de como ocorre na IETF, o processo de construção de padrões é realizado por grupos de trabalhos fechados, sendo restrita a membros, tornando-se mais célere o processo, se comparado com a IETF. Apesar de sua importância, a organização possui pouco engajamento na Governança da *Internet*, como resume Gatto, Moreiras e Getschko: “Embora o W3C seja importante para a *Internet*, a organização não tem despertado muito interesse nos debates em curso sobre a Governança da *Internet*”.²²⁵

- *Internet Assigned Number Authority (IANA)* foi inicialmente criada de maneira informal para executar atividades de gerenciamento dos números de IP, além de gerenciar sistema de domínios. Até o final da década de 90, Jonh Postel²²⁶ era o gestor e coordenador técnico da raiz da *Internet*, executando as funções relativas a IANA²²⁷ até 1998, ano de seu falecimento. Wagner e Canabarro, lembram que Postel “definiu pessoalmente uma série de

²²⁵ GATTO, Raquel F.; MOREIRAS, Antonio M.; e GETSCHKO, Demi. Governança da Internet: conceitos, evolução e abrangência. In. **27º Simpósio Brasileiro de Redes de Computadores e sistemas Distribuídos**, 2009. p. 82.

²²⁶ Foi acadêmico do Instituto de Ciências da Informação da Universidade do Sul da Califórnia. Bem como, “foi responsável por manter, atualizar e distribuir a base de dados da raiz para os administradores dos diversos sistemas autônomos integrantes da Internet; adotou uma série de políticas e práticas relativas à alocação descentralizada de blocos de endereços IP para organizações regionais conhecidas como Registros Regionais (RIRs); e definiu a lógica de funcionamento do sistema de nomes de domínio disponibilizados para entidades públicas e privadas, com ou sem intento lucrativo, que operam no comércio de nomes de domínio que identificam espaços delimitados na Internet e páginas na Web”. (WAGNER, Flávio Rech; CANABARRO, Diego Rafael. A governança da Internet: definição, desafios e perspectivas. In: PIMENTA, Marcelo Soares; CANABARRO, Diego Rafael (Org.). **Governança Digital**. Porto Alegre: Editora da UFRGS/CEGOV, 2014. p. 197).

²²⁷ “As funções IANA referem-se à administração bases de dados centrais que catalogam os inúmeros protocolos da Internet e servem de base para a alocação de endereços numéricos (IPs), bem como para a resolução de nomes de domínio em endereços IPs. Além disso, elas englobam, também, a coordenação do sistema de servidores-raiz da Internet.” (CANABARRO, Diego. Governança Global da Internet: Aspectos Conceituais, Questões da Agenda Contemporânea e Prospectos para o Estudo do Tema. In: OPPERMAN, Daniel. (Ed.). **Internet Governance in the Global South: History, Theory, and Contemporary Debates**. São Paulo: University of São Paulo, 2018. p. 70. Disponível em: <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-65805-2>. Acesso em: 08/07/2022).

políticas que orientavam a normalização de protocolos e a distribuição de números utilizados na rede, bem como a delimitação e o registro de nomes de domínio correspondentes.”²²⁸ Com isso, Jonh Postel foi reconhecido como a *Internet Assigned Numbers Authority* (IANA). Após sua morte, a função de gerenciar e designar os números IP aos integrantes da *Internet* passou a ser executada pelo ICANN. Com isso, a IANA foi incorporada a ICANN, permanecendo com sua função IANA.

Na prática, IANA mesmo detendo a centralização da base do IP, entrega às entidades regionais, as chamadas de *Regional Internet Registries* (RIRs),²²⁹ a construção em conjunto das políticas de distribuição de números de IP.

- *Internet Corporation for Assigned Names and Numbers* (ICANN), foi fundado em 1998 como uma empresa privada ligada ao Departamento de Comércio dos Estados Unidos, por meio de um *Memorandum of Understanding*, sob a égide da legislação do Estado da Califórnia.²³⁰ Passou a ter destaque na governança raiz da *Internet* e considerada por muitos como a principal entidade de Governança da *Internet* da atualidade.²³¹ A ICANN, com a incorporação da IANA, “assumiu o controle político e a coordenação do funcionamento da raiz, bem como a regulação do mercado do DNS,”²³² mantendo consigo o controle matricial da distribuição de identificadores numéricos para sistemas autônomos da Rede.

A ICANN é caracterizada como uma organização plural, assumindo o compromisso, em seu Estatuto, de adotar processos de desenvolvimento de políticas multisetoriais abertos e transparentes, privilegiando a participação do setor privados, levando em consideração a

²²⁸ WAGNER, Flávio Rech; CANABARRO, Diego Rafael. A governança da Internet: definição, desafios e perspectivas. In: PIMENTA, Marcelo Soares; CANABARRO, Diego Rafael (Org.). **Governança Digital**. Porto Alegre: Editora da UFRGS/CEGOV, 2014. p. 197).

²²⁹ “Os RIRs são entidades autoreguladas, que decidem de forma autônoma as políticas de governança sobre os recursos que controlam, nas respectivas regiões. Seus membros são todas as instituições que receberam designações de blocos IP dessas entidades. A atividade dos RIRs, no entanto, tem normalmente ultrapassado o simples gerenciamento dos IPs. Eles promovem atividades educacionais e de disseminação de tecnologias importantes nas respectivas regiões.” (GATTO, Raquel F.; MOREIRAS, Antonio M.; GETSCHKO, Demi. Governança da Internet: conceitos, evolução e abrangência. In. **27º Simpósio Brasileiro de Redes de Computadores e sistemas Distribuídos**, 2009. p. 83).

²³⁰ CAROTTI, Bruno. ICANN and Global Administrative Law. In: CASSESE, Sabino et alli. (eds.). **Global Administrative Law: Cases, Materials, Issues**. New York University School of Law: Institute for International Law and Justice, 2008. p. 3. Disponível em: <https://studylib.net/doc/15758957/icann-and-global-administrative-law--by-bruno-carotti>. Acesso em: 29/05/2022.

²³¹ GATTO, Raquel F.; MOREIRAS, Antonio M.; GETSCHKO, Demi. Governança da Internet: conceitos, evolução e abrangência. In. **27º Simpósio Brasileiro de Redes de Computadores e sistemas Distribuídos**, 2009. p. 83.

²³² WAGNER, Flávio Rech; CANABARRO, Diego Rafael. A governança da Internet: definição, desafios e perspectivas. In: PIMENTA, Marcelo Soares; CANABARRO, Diego Rafael (Org.). **Governança Digital**. Porto Alegre: Editora da UFRGS/CEGOV, 2014. p. 197).

política pública de *aconselhamento* de governos e autoridades públicas.²³³ Assim como a IANA, outro órgão importante integra a ICANN, o comitê de representação de governos (o *Governmental Advisory Comitee – GAC*), responsável pela “elaboração de políticas públicas sobre regulação da *internet* no cenário internacional.”²³⁴ Esse posicionamento adotado da estrutura de “conselheiros” (*advisory*) é alvo de insatisfação e crítica de vários governos, os quais ostentam o posicionamento de assumir posição com poder de decisão dentro da estrutura da ICANN.

Outro ponto importante tem sido a colocação da ICANN como uma instituição internacional,²³⁵ sem, contudo, ter-se constituído legalmente como um organismo internacional, além de outras críticas como: “falta de equilíbrio nos processos de deliberação política que ocorrem no âmbito da organização; dificuldade de se definir os contornos da *accountability* que deve guiar suas ações.”²³⁶

Nesse sentido, Milton Mueller classifica a ICANN, como um “animal institucional”, devido: (a) constituir uma organização cujo assunto de interesse internacional é tutelado por ela e estabelece vínculos com vários países, contudo, não possui um tratado específico definido, formalmente, como uma Organização Internacional; e, (b) apesar de regular assunto

²³³ Conforme previsão da Seção 1.2, (a) (iv), do Estatuto da Internet Corporation para Atribuição de Nomes e Números, com atualização de 02/junho/2022: “[...] (iv): (iv) Empregar processos de desenvolvimento de políticas multissetoriais abertos, transparentes e de baixo para cima, liderados pelo setor privado (incluindo atores empresariais, sociedade civil, comunidade técnica, academia e usuários finais), levando devidamente em conta a política pública aconselhamento de governos e autoridades públicas. Esses processos devem (A) buscar a opinião do público, para cujo benefício a ICANN atuará em todos os eventos, (B) promover decisões bem informadas com base em conselhos de especialistas e (C) garantir que as entidades mais afetadas possam ajudar na política processo de desenvolvimento;” Tradução livre do original: “Employ open, transparent and bottom-up, multistakeholder policy development processes that are led by the private sector (including business stakeholders, civil society, the technical community, academia, and end users), while duly taking into account the public policy advice of governments and public authorities. These processes shall (A) seek input from the public, for whose benefit ICANN in all events shall act, (B) promote well-informed decisions based on expert advice, and (C) ensure that those entities most affected can assist in the policy development process;” (ICANN. **Bylaws for Internet Corporation for Assigned Names and Numbers**. As amended 2 June 2022. Disponível em: <https://www.icann.org/resources/pages/governance/bylaws-en>. Acesso em 20/06/2022.

²³⁴ LIGUORI FILHO, Carlos Augusto. Governança da internet no espaço regulatório global: o Idiossincrático modelo de gestão da ICANN. In: CELLA, José Renato Gaziero; ROVER, Aires Jose; GOMES, Magno Federici. [Coord.]. **Direito, governança e novas tecnologias** [Recurso eletrônico on-line]. Org. CONPEDI/UFGM/FUMEC/Dom Helder Câmara. Florianópolis: CONPEDI, 2015. p. 317.

²³⁵ Como exemplo desse aspecto, possui no Brasil, servidores-raiz espelho operado pelo NIC.br, apesar de características como neutralidade, estabilidade e segurança, integram agenda de vários governos, a concentração num único governo dificulta o gerenciamento na raiz do sistema, caso houvesse algo maculando os interesses da comunidade. A expectativa é a evolução do ICANN para se desvincular do governo estadunidense e adote a forma de organismo internacional. Nesse sentido, Gatto, Moreiras e Getschko revelam: “Espera-se que a ICANN futuramente evolua, assumindo de fato uma posição de organismo internacional, mantendo a natureza de instituição plural e sem fins lucrativos e rompendo o vínculo com o governo estadunidense.” (GATTO, Raquel F.; MOREIRAS, Antonio M.; GETSCHKO, Demi. Governança da Internet: conceitos, evolução e abrangência. In: **27º Simpósio Brasileiro de Redes de Computadores e sistemas Distribuídos**, 2009. p. 85-86).

²³⁶ WAGNER, Flávio Rech; CANABARRO, Diego Rafael. A governança da Internet: definição, desafios e perspectivas. In: PIMENTA, Marcelo Soares; CANABARRO, Diego Rafael (Org.). **Governança Digital**. Porto Alegre: Editora da UFRGS/CEGOV, 2014. p. 197).

que transcende as fronteiras nacionais, ela está diretamente vinculada à legislação de um país específico.²³⁷

Apesar de se atribuir uma relevância demasiada à ICANN, a sua atuação, embora seja importante para o gerenciamento da raiz da *Internet*, em termos de Governança da *Internet*, seu papel é bem mais comedido, restringindo-se mais ao aspecto técnico, levando em consideração tanto seu aspecto legal, como sua função de somente gerenciar a estrutura básica de nomes e a política de distribuição de números IP,²³⁸ que inclusive poderá ser minorada com a instituição do Ipv6 (versão 6 do Protocolo *Internet*).²³⁹

Contudo, destacamos o mecanismo de participação “de baixo para cima” (*bottom-up*),²⁴⁰ o qual permite a ampla participação dos interessados na governança da *Internet*, além de adotar o *consensus driven*,²⁴¹ o processo de decisão por meio de consensos entre os

²³⁷ MUELLER, Milton. **Ruling the Root: Internet Governance and the Taming of Cyberspace**. Cambridge, MA: MIT Press, 2002. p. 217.

²³⁸ GATTO, Raquel F.; MOREIRAS, Antonio M.; GETSCHKO, Demi. Governança da Internet: conceitos, evolução e abrangência. In. **27º Simpósio Brasileiro de Redes de Computadores e sistemas Distribuídos**, 2009. p. 86.

²³⁹ “Com o objetivo de suprir a falta de endereçamento e outras necessidades em relação ao IPv4, na década de 90 foi desenvolvido por um grupo de pesquisadores do Internet Engineering Task Force (IETF) a nova geração do protocolo IP, a next generation, ou Internet Protocol version 6 (IPv6, ou protocolo de internet versão 6), o qual possui espaço de endereçamento de 128 bits, suporte a roteamento e segmentação de pacotes na estação de origem, mecanismos de segurança, entre outros avanços. Em função da necessidade de recursos demandados pelo rápido avanço tecnológico e a utilização de redes convergentes, como stream de video em alta definição e a Internet das Coisas (ou Internet of Things - IoT), a especificação do protocolo IPv6 tem sido revisada continuamente. Com a utilização predominantemente com dispositivos IPv4, se faz necessária a transição para o protocolo IPv6, utilizando recursos como tunelamento e pilha-dupla (dual stack).” (HEY NETTO, Ricardo. **Protocolo de internet versão 6 (IPv6)**. 2018. 53 f. Monografia de Especialização em Redes de Computadores e Teleinformática, Departamento Acadêmico de Eletrônica, Universidade Tecnológica Federal do Paraná. Curitiba, 2018. p. 11-12). Para uma melhor conceituação e compreensão dos fatores econômicos que afetam a transição para o IPv6, conferir: KUEBIS, Brenden; MUELLER, Milton. “The hidden standards war: economic factors affecting IPv6 deployment”, **Digital Policy, Regulation and Governance**. v. 22, N. 4, p. 333-361. 2020. Disponível em: <https://doi.org/10.1108/DPRG-10-2019-0085>. Acesso em: 10/06/2022.

²⁴⁰ Conceito constante no Site da ICANN: “Baixo para cima. Na ICANN, em vez de a Diretoria apenas declarar quais tópicos a ICANN abordará, os membros dos subgrupos da ICANN podem levantar questões em nível de base. Então, se o assunto valer a pena ser abordado e for da alçada da ICANN, ele poderá passar por vários comitês consultivos e organizações de apoio até que as recomendações de políticas sejam encaminhadas à Diretoria para votação.” Tradução livre do original: “Bottom up. At ICANN, rather than the Board of Directors solely declaring what topics ICANN will address, members of sub-groups in ICANN can raise issues at the grassroots level. Then, if the issue is worth addressing and falls within ICANN's remit, it can rise through various Advisory Committees and Supporting Organizations until eventually policy recommendations are passed to the Board for a vote.” (ICANN, disponível em: <https://www.icann.org/resources/pages/welcome-2012-02-25-en>. Acesso em: 16/06/2022).

²⁴¹ “Orientado por consenso. Por meio de seus estatutos, processos e reuniões internacionais, a ICANN oferece a arena onde todos os defensores podem discutir questões de política da Internet. Quase qualquer pessoa pode participar da maioria dos grupos de trabalho voluntários da ICANN, garantindo uma ampla representação das perspectivas do mundo. Ouvir todos os pontos de vista, buscar interesses mútuos e trabalhar em direção ao consenso leva tempo, mas o processo resiste à captura por qualquer interesse único – uma consideração importante ao gerenciar um recurso tão vital quanto a Internet global.” Tradução livre do original: “Consensus-driven. Through its Bylaws, processes, and international meetings, ICANN provides the arena where all advocates can discuss Internet policy issues. Almost anyone can join most of ICANN's volunteer Working Groups, assuring broad representation of the world's perspectives. Hearing all points of view, searching for

participantes, Como forma de destacar a legitimidade e a autoridade da ICANN, com o discurso *online* de que ela é um veículo de expressão da comunidade de *Internet*²⁴², em que “todos os usuários da Internet merecem uma manifestação sobre como ela deve funcionar”²⁴³

- *World Summit on the Information Society* (WSIS) e *Working Group on Internet Governance* (WGIG): analisadas as principais instituições e organismo – sem o ímpeto de realizar um cotejo aprofundado – é possível compreender a importância de cada uma delas na construção de uma governança da *internet* sólida – mesmo diante de tantos desafios, dentre eles, a proteção de direitos fundamentais – representando parte dos atores deste contexto.

Nesse passo, apresentam-se dois atores, forjados a partir de fóruns de debates sobre as novas questões submetidas à regulação nacionais, além do avanço global de tais demandas, as quais não foram tratadas por nenhuma das organizações abordadas anteriormente, além da inegável importância da rede, a União Internacional de Comunicações firmou proposição, resultando na formulação de uma Cúpula Mundial sobre a Sociedade da Informação²⁴⁴ ou em inglês a *World Summit on the Information Society* – WSIS.

A WSIS representa uma cúpula das Nações Unidas, que fora realizada em duas fases, sendo a primeira em 2003, em Genebra, e a segunda em 2005, em Tunis. Com isso, estabeleceu a Governança da *Internet*, como um campo de estudo, proporcionando um ambiente para intensos debates a respeito da definição de governança da *Internet* – tal como

mutual interests, and working toward consensus take time, but the process resists capture by any single interest—an important consideration when managing a resource as vital as the global Internet.” (ICANN, disponível em: <https://www.icann.org/resources/pages/welcome-2012-02-25-en>. Acesso em: 16/06/2022).

242 BIANCHI, José Flávio. Internet e Regulação: a ICANN à luz da teoria da regulação. **Revista de Direito, Estado e Telecomunicações**. Brasília, v. 8, n. 1, p. 135-156, maio 2016. p. 141. Disponível em: <https://doi.org/10.26512/lstr.v8i1.21529>. Acesso em: 16/06/2022.

243 “Modelo multissetorial. A abordagem inclusiva da ICANN trata o setor público, o setor privado e os especialistas técnicos como pares. Na comunidade da ICANN, você encontrará registros, registradores, provedores de serviços de Internet (ISPs), defensores da propriedade intelectual, interesses comerciais e empresariais, interesses não comerciais e sem fins lucrativos, representação de mais de 100 governos e uma variedade global de Usuários de internet. Todos os pontos de vista são considerados por seus próprios méritos. A crença fundamental da ICANN é que todos os usuários da Internet merecem uma opinião sobre como ela é administrada.” Tradução livre do original: “Multistakeholder model. ICANN's inclusive approach treats the public sector, the private sector, and technical experts as peers. In the ICANN community, you'll find registries, registrars, Internet Service Providers (ISPs), intellectual property advocates, commercial and business interests, non commercial and non-profit interests, representation from more than 100 governments, and a global array of individual Internet users. All points of view receive consideration on their own merits. ICANN's fundamental belief is that all users of the Internet deserve a say in how it is run.” (ICANN, disponível em: <https://www.icann.org/resources/pages/welcome-2012-02-25-en>. Acesso em: 16/06/2022).

244 Nesse sentido, resume Diego Canabarro: “A Cúpula começou a ser gestada no final da década de 1990, quando a União Internacional das Telecomunicações (UIT) comissionou ao Secretário-Geral da ONU a realização de um processo continuado destinado a gerar insumos relacionados ao papel que a “sociedade da informação” poderiam ter para o alcance das Metas de Desenvolvimento do Milênio.” (CANABARRO, Diego. Governança Global da Internet: Aspectos Conceituais, Questões da Agenda Contemporânea e Prospectos para o Estudo do Tema. In: OPPERMANN, Daniel. (Ed.). **Internet Governance in the Global South: History, Theory, and Contemporary Debates**. São Paulo: University of São Paulo, 2018. p. 79. Disponível em: <https://nbn-resolving.org/urn:nbn:de:0168-ssor-65805-2>. Acesso em: 08/07/2022).

descrito acima – havendo posições bem distintas, desde do gerenciamento de recursos críticos da *Internet* pela ICANN, até a regulação de todo o espectro da Tecnologia de Informação e Comunicação (TIC).²⁴⁵

Além da construção de uma definição de governança da *Internet*, e sua constituição como um campo de estudo, a WSIS inseriu oficialmente a temática de governança da *Internet* nas pautas diplomáticas. Dentre outros resultados nesse assunto, a Cúpula estabeleceu um Grupo de Trabalho da *Internet* na ONU, o denominado *Working Group on Internet Governance* (WGIG).

A criação efetiva do WGIG ocorreu no ano de 2004, sendo seus membros indicados pela ONU, seguindo uma lista de membros integrantes de diversas áreas, como: governos, agências internacionais e multilaterais, sociedade civil e a iniciativa privada. A principal missão do grupo foi concatenar toda discussão sobre a governança da *Internet* em temas centrais, demandando uma atuação rápida e urgente, a saber: 1. A supervisão unilateral do arquivo de zona raiz pelo governo estadunidense; 2. Os custos de interconexão; 3. Estabilidade da Rede, segurança e crime cibernético; 4. Spam; 5. Participação na formulação global de políticas; 6. Construção de capacidades; 7. Alocação de nomes de domínio; 8. Endereçamento IP; 9. Direitos de propriedade intelectual; 10. Liberdade de expressão; 11. Proteção de dados pessoais; 12. Proteção ao consumidor; 13. Diversidade linguística.²⁴⁶

Outra contribuição decisiva da WGIG, como lembram Musiani e Schafer²⁴⁷, foi nos esforços de definição da governança da *internet*, pontuando sobre “princípios, normas, regras, procedimentos de tomada de decisão e programas compartilhados que moldam a evolução e o uso da *Internet*”, além do mais, estipulou que a governança da *Internet* constitui uma demanda denominada de “multissetorial”,²⁴⁸ com atores relevantes não apenas Estados-nação,

²⁴⁵ MUSIANI, Francesca; SCHAFER, Valérie. Global Governance: A Short History of Debates Born With the Telegraph and Popularized by the Internet. In: BALBI, Gabriele; RIBEIRO, Nelson; SCHAFER, Valérie; SCHWARZENEGGER, Christian. [Org.]. **Digital Roots: Historicizing Media and Communication Concepts of the Digital Age**. Berlin, Boston: De Gruyter Oldenbourg, 2021. p. 122, Disponível em: <https://doi.org/10.1515/9783110740202-007>. Acesso em: 25/10/2021.

²⁴⁶ NAÇÕES UNIDAS (2005). Report of the Working Group on Internet Governance. Disponível em: <http://www.wgig.org/docs/WGIGREPORT.pdf>. Acesso em: 22/06/2022.

²⁴⁷ MUSIANI, Francesca; SCHAFER, Valérie. Global Governance: A Short History of Debates Born With the Telegraph and Popularized by the Internet. In: BALBI, Gabriele; RIBEIRO, Nelson; SCHAFER, Valérie; SCHWARZENEGGER, Christian. [Org.]. **Digital Roots: Historicizing Media and Communication Concepts of the Digital Age**. Berlin, Boston: De Gruyter Oldenbourg, 2021. p. 122, Disponível em: <https://doi.org/10.1515/9783110740202-007>. Acesso em: 25/10/2021.

²⁴⁸ O multissetorialismo constitui um novo arranjo de governança, que se tornou o principal assunto de pesquisa sobre governança da Internet para acadêmicos, em diversos campos. Conferir a respeito: BRADSHAW, Samantha; DENARDIS, Laura; HAMPSON, Fen; JARDINE, Eric; RAYMOND, Mark. **The Emergence of Contention in Global Internet Governance**. 56ª Convenção Anual da ISA. New Orleans, USA. 2015. Disponível em: <http://web.isanet.org/Web/Conferences/New%20Orleans%202015/Archive/adf6039d-3251-4ede-a2fa-0f8701f21459.pdf>. Acesso em: 25/06/2022.

mas também a sociedade civil em suas diferentes facetas (comunidade técnica, associações de liberdades civis e cidadãos em sua capacidade), além de empresas.

Musiani e Schafer, pontuam haver certa frustração, devido alguns autores desejarem a suplantação, por parte da WSIS, do papel, até então, unilateral e predominante dos Estados Unidos na ICANN, contudo não se conseguiu atingir tal fato.²⁴⁹ Porém, um dos maiores resultados da Conferência de Tunis, foi a origem de um espaço global de discussão sobre governança da *Internet*, o qual continua até hoje: o *Internet Governance Forum* (IGF).

- O *Internet Governance Forum* (IGF), sob responsabilidade do Secretário-Geral das Nações Unidas, constitui o principal fórum global multilateral para debater políticas públicas voltadas à governança da *Internet*, desde a primeira reunião em 2006, em Atenas²⁵⁰, tendo como objetivo “[...] ser uma plataforma de diálogo e intercâmbio de pontos de vista multissetorial, imparcial e independente, e de compartilhamento de conhecimentos e melhores práticas sobre políticas relativas à *Internet*” [...] ²⁵¹, constitui um fórum aberto a todas as pessoas com interesse em questões de governança da *Internet*.

Conforme consta nos parágrafos 72 a 78 da Agenda de Túnis²⁵², as reuniões do IGF devem ocorrer periodicamente a cada cinco anos, porém, tal prazo, inicialmente acordado na Agenda, pode ser prorrogado indefinidamente por decisão do Secretário Geral das Nações Unidas, devendo ser composto por representantes de governos, sociedade civil, iniciativa privada e organizações internacionais, cabendo ao Secretário Geral convocar os participantes e zelar pelo equilíbrio de representação geográfica na reunião.

Ademais, o Fórum não possui pretensões de supervisionar ou, até mesmo, substituir arranjos, instituições ou organizações já existentes. Sua colaboração, cinge em gerenciar, juntamente com esses atores, soluções para os desafios da governança, auxiliando na

²⁴⁹ MUSIANI, Francesca; SCHAFFER, Valérie. Global Governance: A Short History of Debates Born With the Telegraph and Popularized by the Internet. In: BALBI, Gabriele; RIBEIRO, Nelson; SCHAFFER, Valérie; SCHWARZENEGGER, Christian. [Org.]. **Digital Roots: Historicizing Media and Communication Concepts of the Digital Age**. Berlin, Boston: De Gruyter Oldenbourg, 2021. p. 122, Disponível em: <https://doi.org/10.1515/9783110740202-007>. Acesso em: 25/10/2021.

²⁵⁰ Já foram realizadas 15 (quinze) reuniões anuais do IGF, sediadas por diversos governos. A reunião anual do IGF de 2021 foi realizada pelo Governo da Polônia, em Katowice. Conforme o planejamento atual, o IGF de 2022 será organizado pelo Governo da Etiópia, o IGF de 2023, pelo Governo do Japão e o IGF de 2025, pela Federação Russa. O anfitrião do IGF de 2024 ainda não foi confirmado, mas o Governo do Canadá já se candidatou para sediar a reunião em Montreal.

²⁵¹ NAÇÕES UNIDAS. **O Fórum de Governança da Internet**. Disponível em: <https://brasil.un.org/pt-br/71321-o-que-e-o-forum-de-governanca-da-internet-igf>. Acesso em: 25/06/2022.

²⁵² World Summit on the Information Society. **Tunis agenda for the information Society**. 2005. Disponível em: <https://www.itu.int/net/wsis/docs2/tunis/off/6rev1.html>. Acesso: 15/05/2022.

promoção de debates e incentivando cada vez mais a participação no processo, sem se envolver no cotidiano dos demais atores, ou nas operações técnicas da *Internet*.²⁵³

Até aqui, demonstramos um mapeamento de organizações componentes do universo da governança da *Internet* – logicamente não é objeto deste estudo exaurir a análise envolvendo cada uma delas – contudo o panorama apresentado aclara, resumidamente, o desenvolvimento da governança da *Internet*, bem como busca solidificar a noção múltipla de atores com influência no discurso *online*, por meio de suas atividades.

Nesse contexto, inclui a atuação de organizações internacionais, em meio ao cenário de mudanças, ultrapassando as margens do processo – de meramente coordenar entre estados-nação – para o núcleo, interagindo com questões de governança da *Internet*.²⁵⁴, além do Estado-nação, por meio das legislações e medidas adotadas institucionalmente que afetam as operações dentro do território. Por fim, as empresas privadas que exploram serviços *online*, participando do debate em fóruns internacionais, bem como desenvolvendo e praticando suas próprias normas de governança privada – principalmente as empresas de mídias sociais.

Neste último ponto, consiste a preocupação deste estudo, considerando a capacidade em livre evolução das empresas de mídia social na vigilância (algorítmica) e formulação de controle privado no desenvolvimento de suas normas próprias, criando uma forma de governança privada. Como explica Balkin, em sua visão triangular da liberdade de expressão

²⁵³ Resumindo esse pensamento, conferir o texto do Parágrafo 77, da Agenda de Túnis: “O IGF não teria função de supervisão e não substituiria os arranjos, mecanismos, instituições ou organizações existentes, mas os envolveria e tiraria proveito de sua experiência. Seria constituído como um processo neutro, não duplicativo e não vinculativo. Não teria nenhum envolvimento no dia-a-dia ou nas operações técnicas da Internet”. Tradução livre do original: “77. The IGF would have no oversight function and would not replace existing arrangements, mechanisms, institutions or organizations, but would involve them and take advantage of their expertise. It would be constituted as a neutral, non-duplicative and non-binding process. It would have no involvement in day-to-day or technical operations of the Internet.” World Summit on the Information Society. **Tunis agenda for the information Society**. 2005. Disponível em: <https://www.itu.int/net/wsis/docs2/tunis/off/6rev1.html>. Acesso: 15/05/2022.

²⁵⁴ Nesse sentido Levison e Marzouki, discorrem: “Esse modelo não desapareceu e as OIs com suas secretarias ainda desempenham esse papel ou, no mínimo, dão um discurso a esse papel. No entanto, em meio a contextos em mudança, incertezas crescentes, tanto técnicas quanto geopolíticas, e uso crescente do termo, bem como a presença de ‘multissetoriais’, as OIs reformularam ativamente seus papéis na promulgação de ideias de participação multissetorial no ecossistema de governança da Internet, bem como na reformulação de modelos de colaboração e conflito por meio de ligações informais e formais com outros grupos de partes interessadas, definindo agendas de pesquisa, promovendo o ‘diálogo’ e até organizando conferências.” Tradução livre do original: “This model has not disappeared and IOs with their secretariats still do play this role or, at the very least, give lip service to this role. Yet amidst changing contexts, increasing uncertainties both technical and geopolitical, and growing usage of the term as well as presence of ‘multistakeholders’, IOs have actively reshaped their roles in promulgating ideas of multistakeholder participation in the internet governance ecosystem as well as reshaping models of collaboration and conflict through informal and formal linkages to other stakeholder groups, defining research agendas, fostering ‘dialogue’, and even hosting conferences.” LEVINSON, Nanette S.; MARZOUKI, Meryem. *Global Internet Governance and Development: International Organizations Roles. A 112^a Annual Meeting, American Political Science Association - APSA*. Setembro de 2017, Philadelphia (PA), United States. Disponível em: <https://hal.archives-ouvertes.fr/hal-01300442>. Acesso em: 25/10/2021.

– inclusive amoldada na definição multissetorial de governança – essas empresas detentoras de uma imensa infraestrutura privada, devem ser encaradas não apenas como operadores, facilitando ou dificultando a comunicação *online*, mas como governantes do discurso *online*.²⁵⁵ Ou seja, “em vez de ver as empresas de infraestrutura digital como meros canais ou plataformas, devemos reconhecê-las como governantes dos espaços sociais.”²⁵⁶

Realizada a compreensão da dimensão subjetiva, passamos a demonstrar a dimensão objetiva, a qual guarda peculiar importância para o trabalho, ao completar a linha de compreensão: a moderação de conteúdo como prática de governança privada.

2.3. Governança (privada) da *internet*: dimensão objetiva

A dimensão objetiva trata de como a governança da *Internet* se materializa, por meio de suas diversas práticas, dentre elas, a coordenação de nome de domínios realizadas pelas pelo ICANN; na elaboração de leis nacionais objetivando a neutralidade de rede, dispondo sobre a não discriminação de conteúdo por provedores de acesso; pela criação de filtro de conteúdo impróprios para crianças, por provedores de serviços; além da exclusão de conteúdo e, ou, contas de usuários, considerados impróprios por determinada rede social. Essas práticas, possuem inegável efeito na operação da *Internet*, não só na visão de quem gerencia, mas também, de quem utiliza o serviço, devendo, portanto, integrar a governança.

A despeito dessa constatação, como ficou evidenciado em matéria de governança da *Internet*, não há uma direção concisa e uniforme, ao contrário, por envolver diversas áreas, há formas distintas de abordagem e materialização, influenciando no sistema de governança. Nesse sentido, Kurbalija, na introdução de seu livro “*An Introduction to Internet Governance*”, resume:

A forma como a Internet é definida reflete diferentes perspectivas, abordagens e interesses políticos. Normalmente, os especialistas em telecomunicações veem a governança da Internet sob o prisma do desenvolvimento de uma infraestrutura técnica. Especialistas em computação se concentram no desenvolvimento de diferentes padrões e aplicações, como XML (eXtensible Markup Language) ou Java. Especialistas em comunicação enfatizam a facilitação da comunicação. Ativistas de direitos humanos veem a governança da Internet da perspectiva da liberdade de

²⁵⁵ BALKIN, Jack M. Free Speech is a triangle. **Columbia Law Review**. v. 118, p. 2011–2055, 2018, p. 2021. Disponível em: <https://openyls.law.yale.edu/bitstream/handle/20.500.13051/17953/Balkin%2C%20Free%20Speech%20is%20a%20Triangle.pdf?sequence=1&isAllowed=y>. Acesso em: 04/11/2021.

²⁵⁶ BALKIN, Jack M. Free Speech is a triangle. **Columbia Law Review**. v. 118, p. 2011–2055, 2018, p. 2021. Disponível em: <https://openyls.law.yale.edu/bitstream/handle/20.500.13051/17953/Balkin%2C%20Free%20Speech%20is%20a%20Triangle.pdf?sequence=1&isAllowed=y>. Acesso em: 04/11/2021.

expressão, privacidade e outros direitos humanos básicos. Os advogados concentram-se na jurisdição e na resolução de disputas. Políticos em todo o mundo geralmente se concentram em questões que repercutem em seus eleitores, como tecno-otimismo (mais computadores = mais educação) e ameaças (segurança na Internet, proteção infantil). Os diplomatas estão principalmente preocupados com o processo e a proteção dos interesses nacionais. A lista de perspectivas profissionais potencialmente conflitantes da governança da Internet continua.²⁵⁷

Nesse contexto multidisciplinar, seguindo a tese baseada no conceito de camadas²⁵⁸ como característica fundamental da *Internet*, proposta por Solum e Chung, destaca-se o fato de as camadas serem organizadas em uma hierarquia vertical, iniciando pela “camada de conteúdo (o nível “mais alto”) através das camadas de aplicação, transporte, IP e *link* para a camada física (o nível “mais baixo”); por meio da camada física em pacotes; e, em seguida, flui de volta através das mesmas camadas na ordem inversa.”²⁵⁹ Os autores defendem essa tese, formulando o que chamam de “princípio das camadas”, como regra geral para os reguladores da *Internet*: respeitar a integridade das camadas. Assumindo dois postulados: “o princípio da separação de camadas: a regulamentação da *Internet* não deve violar ou comprometer a separação entre as camadas projetadas na arquitetura básica da *Internet*.”²⁶⁰; e o segundo, sendo o “princípio da minimização do cruzamento de camadas: minimizar a distância entre a camada na qual a lei visa produzir efeito e a camada diretamente afetada pela regulamentação legal.”²⁶¹

²⁵⁷ KURBALIJA, Jovan. **An Introduction to Internet Governance**. 7 ed. Geneva: Diplo Foundation, 2016, p. 5. Disponível em: https://www.diplomacy.edu/wp-content/uploads/2021/12/AnIntroductiontoIG_7th-edition.pdf. Acesso em: 20/06/2022.

²⁵⁸ Segundo os autores, o sistema de comunicação entre usuários, é composta por seis camadas que constituem a Internet, a saber: “A Camada de Conteúdo: Os símbolos e imagens que são comunicados; A Camada de Aplicação: Os programas que usam a Internet, por exemplo, a Web; A camada de transporte: TCP, que divide os dados em pacotes; A Camada de Protocolo da Internet: IP, que trata do fluxo de dados pela rede; A Camada de Link: A interface entre os computadores dos usuários e a camada física; e A Camada Física: O fio de cobre, o cabo óptico, os links de satélite, etc. Apresentamos esta descrição esquelética com mais detalhes abaixo. (Grifouse). Tradução livre do original: “The Content Layer: The symbols and images that are communicated; The Application Layer: The programs that use the Internet, e.g., the Web; The Transport Layer: TCP, which breaks the data into packets; The Internet Protocol Layer: IP, which handles the flow of data over the network; The Link Layer: The interface between users' computers and the physical layer; and The Physical Layer: The copper wire, optical cable, satellite links, etc. We flesh out this skeletal description in greater detail below.” (SOLUM, Lawrence B.; CHUNG, Minn. The Layers Principle: Internet Architecture and the Law. **Notre Dame Law Review**. v. 79, N. 03, p. 815-948. 2004. p. 816. Disponível em: <http://scholarship.law.nd.edu/ndlr/vol79/iss3/1>. Acesso em: 22/06/2022).

²⁵⁹ SOLUM, Lawrence B.; CHUNG, Minn. The Layers Principle: Internet Architecture and the Law. **Notre Dame Law Review**. v. 79, N. 03, p. 815-948. 2004. p. 816. Disponível em: <http://scholarship.law.nd.edu/ndlr/vol79/iss3/1>. Acesso em: 22/06/2022.

²⁶⁰ Tradução livre do original: “[...] *the principle of layer separation: Internet regulation should not violate or compromise the separation between layers designed into the basic architecture of the Internet.*” (SOLUM, Lawrence B.; CHUNG, Minn. The Layers Principle: Internet Architecture and the Law. **Notre Dame Law Review**. v. 79, N. 03, p. 815-948. 2004. p. 817. Disponível em: <http://scholarship.law.nd.edu/ndlr/vol79/iss3/1>. Acesso em: 22/06/2022).

²⁶¹ Tradução livre do original: “[...] *is the principle of minimizing layer crossing: minimize the distance between the layer at which the law aims to produce an effect and the layer directly affected by legal regulation.*” SOLUM, Lawrence B.; CHUNG, Minn. The Layers Principle: Internet Architecture and the Law. **Notre Dame**

Diferentemente dessa construção, o estudo de DeNardis e Musiani opta por examinar a governança *pela* infraestrutura da *Internet*, em vez da governança *da* infraestrutura da *Internet*, denominado por elas de “*turn to infrastructure in Internet governance*”²⁶² – em português, numa tradução livre, a “virada para a infraestrutura na governança da *Internet*” – e lançam quatro parâmetros para analisar as infraestruturas de governança da *Internet*, conforme esse prisma, devendo-se considerar:

Primeiro, o estudo de como a Internet é governada é diferente do estudo de como as pessoas usam a Internet. Em segundo lugar, o estudo da governança da Internet está relacionado ao conjunto de tecnologias específicas da Internet, em vez de infraestruturas de tecnologia de informação e comunicação mais amplas (por exemplo, políticas relacionadas a espectro ou microprocessadores). Terceiro, esses exames devem se estender além das instituições (por exemplo, ICANN) para abranger como a governança também é promulgada por meio de design técnico, políticas do setor privado e leis nacionais e tratados internacionais. Finalmente, a governança da Internet, infelizmente, não inclui apenas esforços para manter a Internet segura, livre e funcionando eficientemente, mas também abrange abordagens de governança projetadas para restringir o fluxo de informações ou decretar sistemas de vigilância e censura.²⁶³

Nesse panorama, fixamos como importante para este trabalho as práticas de governança desenvolvidas na camada expositiva da *Internet*, sobretudo quanto à moderação do conteúdo realizada pelas empresas de mídias sociais em seus modelos de negócios, a qual observa os parâmetros determinados na legislação nacional e em tratados internacionais àquelas situações. Ratificando, como adverte Keller, que “a atividade reguladora não exaure

Law Review. v. 79, N. 03, p. 815-948. 2004. p. 817. Disponível em: <http://scholarship.law.nd.edu/ndlr/vol79/iss3/1>. Acesso em: 22/06/2022).

²⁶² Isso porque, segundo as autoras: “[...]. Mas em uma era em que as leis nacionais em relação ao conteúdo não se adaptam mais perfeitamente à arquitetura globalmente dispersa e descentralizada da Internet global, há um crescente reconhecimento de que pontos de controle de infraestrutura podem servir como proxies para recuperar (ou obter) controle ou manipular o fluxo de dinheiro, informações e o mercado de ideias na esfera digital.” Tradução livre do original: “But in an era in which nation-bound laws regarding content no longer neatly comport with the globally dispersed and decentralized architecture of the global Internet, there is increasing recognition that points of infrastructural control can serve as proxies to regain (or gain) control or manipulate the flow of money, information, and the marketplace of ideas in the digital sphere.” (DENARDIS, Laura, MUSIANI, Francesca. *Governance by Infrastructure: Introduction*. In MUSIANI, F., COGBURN, Derrick L., DENARDIS, Laura, LEVINSON, Nannete S. (Orgs.) **The Turn to Infrastructure in Internet Governance. Information Technology and Global Governance**. Palgrave Macmillan, New York 2016. Disponível em: <https://ssrn.com/abstract=2730689> or <http://dx.doi.org/10.2139/ssrn.2730689>. Acesso em: 16/06/2022).

²⁶³ Tradução livre do original: “First, the study of how the Internet is governed is distinct from the study of how people use the Internet. Second, the study of Internet governance relates to the set of technologies specific to the Internet rather than broader information and communication technology infrastructures (e.g. policies regarding spectrum or microprocessors). Third, those examinations must extend beyond institutions (e.g. ICANN) to encompass how governance is also enacted through technical design, private industry policies, and national laws and international treaties. Finally, Internet governance, unfortunately, not only includes efforts to keep the Internet secure, free, and efficiently functioning, but also encompasses governance approaches designed to restrict the flow of information or enact systems of surveillance and censorship.” (DENARDIS, Laura, MUSIANI, Francesca. *Governance by Infrastructure: Introduction*. In MUSIANI, F., COGBURN, Derrick L., DENARDIS, Laura, LEVINSON, Nannete S. (Orgs.) **The Turn to Infrastructure in Internet Governance. Information Technology and Global Governance**. Palgrave Macmillan, New York 2016. Disponível em: <https://ssrn.com/abstract=2730689> or <http://dx.doi.org/10.2139/ssrn.2730689>. Acesso em: 16/06/2022).

as formas como os Estados-nação influenciam essa rede de governança”.²⁶⁴ A moderação de conteúdo será estudada, com mais vagar, no próximo capítulo, no qual serão abordadas as principais habilidades tecnológicas e operacionais em relação à regulação e controle dos discursos pelas principais redes sociais, dentre elas o *Facebook*.

2.4. Governança das plataformas²⁶⁵ e direitos humanos: o tom constitucional dos termos de uso

É facilmente perceptível a representação dos avanços proporcionados pelas redes sociais, citando-se, inclusive, no aspecto democrático. A movimentação em massa gerada, segundo o seu modelo de negócio, prevê uma quantidade crescente de usuários nas plataformas, por diversos motivos, trazendo à tona, às empresas privadas, a concretização de uma sociedade *online*. Partindo desta premissa de sociedade virtual, tem-se a analogia das plataformas e suas normas reguladoras serem vistas como Estado e legislação, respectivamente, em prol dos seus desideratos, os quais não estão, necessariamente, associados ao bem-comum.

Consoante a esses acontecimentos, houve a mobilização de usuários contra o *Facebook*, em de fevereiro de 2009, após a notícia da mudança de seus termos de uso, não mais permitindo aos usuários a retirada de seus conteúdos da plataforma a qualquer momento e autorizando, automaticamente, ao *Facebook*, o direito de usar o conteúdo por tempo indeterminado, o que desencadeou uma intensa e abrupta reação do público, pois, em apenas dois dias, mais 100.000 usuários da rede social juntaram-se em protesto, dentro da mesma plataforma, contra essa alteração, forçando Mark Zuckerberg, fundador do *Facebook*, recuar e anunciar a restauração dos termos anteriores, ao mesmo tempo em que seria desenvolvido um trabalho sobre as principais preocupações apontadas pelos usuários sobre os termos de uso, criando, para tanto, um grupo denominado Declaração de Direitos e Responsabilidades do *Facebook*.

²⁶⁴ KELLER, Clara Iglesias. **Regulação nacional de serviços na internet: exceção, legitimidade e o papel do Estado**. Rio de Janeiro: Lumen Juris, 2019. p. 91.

²⁶⁵ Termo cunhado por Robert Gorwa, “um conceito destinado a capturar as camadas de relacionamentos de governança que estruturam as interações entre as partes-chave na sociedade de plataforma de hoje, incluindo empresas de plataforma, usuários, anunciantes, governos e outros atores políticos.” Tradução livre do original: “[...] a concept intended to capture the layers of governance relationships structuring interactions between key parties in today's platform society, including platform companies, users, advertisers, governments, and other political actors.” (GORWA, Robert. What is platform governance? **Information, Communication & Society**. v. 22, n. 6, p. 854-871, 2019. Disponível em: 10.1080/1369118X.2019.1573914. Acesso em: 19/05/2022.

O mais emblemático neste fato é o pronunciamento de Zuckerberg, conforme noticiou o *Electronic Frontier Foundation*, exaltando a importância da participação da comunidade na construção de melhores termos de uso, confira:

Mais de 175 milhões de pessoas usam o *Facebook*. Se fosse um país, seria o sexto país mais populoso do mundo. Nossos termos não são apenas um documento que protege nossos direitos; é o documento que rege como o serviço é usado por todos em todo o mundo. Dada a sua importância, precisamos garantir que os termos reflitam os princípios e valores das pessoas que usam o serviço.²⁶⁶

Nas entrelinhas do discurso, verifica-se, que a ideia de plataforma digital estática, como mera atuação intermediária, não mais persiste e é possível constatar companhias – e não meras plataformas – detentoras de poder político, econômico e social, com dominância de mercado. A crescente economia digital demonstra a tendência ao monopólio na natureza das plataformas digitais. Conforme apontou o relatório produzido pelo *Stigler Committee on Digital Platforms*, o “mercado digital apresenta diversas características que, embora não sejam novas, nunca haviam aparecidos juntas dessa forma, e que empurram o mercado na direção do monopólio”.²⁶⁷

Desta forma, a crítica de Francisco Balaguer Callejón²⁶⁸ disserta sobre a influência das companhias de redes sociais no desenvolvimento democrático, elencando, em síntese, (i) como um dos problemas, não necessariamente derivado das redes sociais em si, mas da atividade das companhias tecnológicas que gerenciam esses aparatos, além do modelo de negócio no qual se baseiam – consistente na coleta de dados dos usuários, indiscriminadamente, fornecendo um grande poder de mercado a poucas companhias, promovendo ou permitindo as instabilidades políticas, a radicalização do espaço público como forma de garantir receita à publicidade, no lugar de manutenção de um clima de estabilidade e consenso. Aponta, (ii) como outro problema, a atividade de agentes de diferentes companhias tecnológicas, os quais utilizam as próprias redes para privilegiar os interesses econômicos das empresas as quais gerenciam as redes. Por fim, (iii) indica o terceiro problema estando vinculado ao tipo de suporte e no desenvolvimento tecnológico, o qual possui uma reduzida capacidade de intervenção pública sobre elas. Concluindo ser

²⁶⁶ HOFMANN, Marcia. Facebook Battle Ends in Major Victory for Users, But the War Continues. **Electronic Frontier Foundation**. 2009. Disponível em: <https://www.eff.org/uk/deeplinks/2009/02/facebook-battle-ends-major-victory-users>. Acesso em: 20/06/2022.

²⁶⁷ LEITÃO, Clara. Plataformas digitais e regulação da neutralidade da rede: como a regulação atende aos interesses de companhias com dominância de mercado. **Revista de Direito, Governança e Novas Tecnologias**. v. 7, n. 1, p. 23-39, 2021. p. 27.

²⁶⁸ CALLEJÓN, Francisco Balaguer. Redes Sociais, Companhias Tecnológicas e Democracia. **REI - Revista Estudos Institucionais**, [S.l.], v. 6, n. 2, p. 579-599, set. 2020. p. 583-584. Disponível em: <https://www.estudosinstitucionais.com/REI/article/view/485> ou <https://doi.org/10.21783/rei.v6i2.485>. Acesso em: 26/06/2022.

“fundamental estabelecer os limites reais, por que estas companhias tendem a utilizar a tecnologia como um obstáculo permanente frente a qualquer controle, afirmando habitualmente haver limitações técnicas que impedem que estes controles sejam implementados.”²⁶⁹

Por essa perspectiva de proeminência das empresas de redes sociais, acrescentando-se as diversas discussões presentes em diferentes camadas cognitivas, o discurso apresentado pelo CEO do *Facebook* menciona a função exploratória das empresas das redes sociais, as quais figuram com equiparação de funções públicas. Como se extrai da metáfora utilizada por Zuckerberg, o qual eleva *Facebook* a um grande Estado: entidade com território próprio – a plataforma digital; sua própria população – os usuários; e por fim, e mais importante, sua própria lei – os termos de serviço.²⁷⁰

Com essa retórica, examinamos a ideia desenvolvida por Edoardo Celeste, a qual denominou o “*tom constitucional*” aos termos de serviços elaborados pelas plataformas de mídias sociais²⁷¹ – especialmente do *Facebook*. Ato contínuo, o delineado acima (item 2.3), faz referência às empresas de redes sociais assumindo o papel representativo de um dos atores que compõe a governança da *Internet*, criando e estipulando suas próprias normas de acordo com seus princípios, revelando um novo método de governança privada.²⁷²

Conforme o supracitado, Balkin, já apontava em seus estudos as empresas de mídias sociais, frente ao modelo de negócio e a crescente popularidade das redes sociais, contrapondo todos os tipos de usuários pelo mundo – além de agentes de estados-nação – e, com isso, fomentaram a ideia de precisar governar suas comunidades, ou seja, elaborar suas

²⁶⁹ CALLEJÓN, Francisco Balaguer. Redes Sociais, Companhias Tecnológicas e Democracia. **REI - Revista Estudos Institucionais**, [S.l.], v. 6, n. 2, p. 579-599, set. 2020. p. 584. Disponível em: <https://www.estudosinstitucionais.com/REI/article/view/485> ou <https://doi.org/10.21783/rei.v6i2.485>. Acesso em: 26/06/2022.

²⁷⁰ CELESTE, Edoardo. Terms of Service and Bill of Rights: New Mechanisms of Constitutionalisation in the Social Media Environment? **Internacional Review of Law, Computers & Technology**. v. 33, n. 02, p. 122-138. 2019. Disponível em: <https://doi.org/10.1080/13600869.2018.1475898>. Acesso em 26/06/2022.

²⁷¹ CELESTE, Edoardo. Terms of Service and Bill of Rights: New Mechanisms of Constitutionalisation in the Social Media Environment? **Internacional Review of Law, Computers & Technology**. v. 33, n. 02, p. 122-138. 2019. Disponível em: <https://doi.org/10.1080/13600869.2018.1475898>. Acesso em 26/06/2022.

²⁷² Nessa perspectiva, o Relatório do Relator Especial sobre a promoção e proteção do direito à liberdade de opinião e expressão, Frank La Rue – do Conselho de Direitos Humanos das Nações Unidas, acentua a participação dos atores privados, na facilitação do desenvolvimento novas dinâmicas de comunicação. Contudo, nesse desenvolvimento é preciso que os atores corporativos adotem medidas que proporcione a vigilância do Estado nas comunicações. Devendo evitar a implantação de tecnologias que permitam a vigilância em massa ou invasiva violando as normas legais vigentes, bem como a observância dos direitos humanos, exigindo que respeitem e promovam os direitos à liberdade de expressão e privacidade, além da proteção dos indivíduos de violações de direitos humanos perpetradas por atores corporativos. (UNITED NATIONS HUMAN RIGHTS COUNCIL. **Report 23/40 of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression**. UN. New York, 17 April 2013. A/HRC/23/40. p. 19-20. Disponível em: <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G13/133/03/PDF/G1313303.pdf?OpenElement>. Acesso em: 11/12/2021).

normas e princípios, fazendo valer e exigindo seu cumprimento no âmbito de suas comunidades *online*.²⁷³ Justificando, sua atuação também a partir da necessidade de construir e manter um espaço aprazível aos usuários:

Os proprietários de plataformas impõem essas regras por motivos mistos. Mas uma razão pela qual fazem isso é fazer com que as pessoas se sintam seguras e respeitadas nas comunidades online. Se os usuários finais se sentirem seguros, eles continuarão a participar, postar conteúdo e tornar a plataforma parte de suas vidas diárias. Isso permite que os proprietários de plataformas sejam lucrativos. Mas também ajuda a promover um fluxo constante e vibrante de ideias e opiniões.²⁷⁴

Percebeu-se, com o tempo, as empresas de mídia social, as quais, em sua concepção, desenvolviam suas atividades tão somente como empresas de tecnologia, iniciando a sua atuação como “governadores comunitários e desenvolveram burocracias elaboradas, que são efetivamente estruturas de governança”.²⁷⁵

Essa visão das plataformas como governadores também é endossada pela professora Kate Klonick, em trabalho seminal, sobre o surgimento de mecanismos internos em empresas de mídia social, a exemplo de *Facebook*, *Twitter* e *YouTube*, em breve síntese, afirma:

[...] as plataformas devem ser consideradas operando como os Novos Governadores do discurso online. Esses Novos Governadores fazem parte de um novo modelo triádico de fala que fica entre o Estado e as editoras dos palestrantes. São entidades privadas, autorreguladoras, motivadas econômica e normativamente para refletir a cultura democrática e as expectativas de liberdade de expressão de seus usuários.²⁷⁶

Klonick, explica o conceito de governança comunitária e a criação de grandes arranjos globais surgindo, praticamente, de forma acidental, sem um estudo prévio, pois as empresas

²⁷³ BALKIN, Jack M. Free Speech is a triangle. **Columbia Law Review**. v. 118, p. 2011–2055, 2018. p. 2021. Disponível em: <https://openyls.law.yale.edu/bitstream/handle/20.500.13051/17953/Balkin%2C%20Free%20Speech%20is%20a%20Triangle.pdf?sequence=1&isAllowed=y>. Acesso em: 04/11/2021.

²⁷⁴ Tradução livre do original: “Platform owners impose these rules out of mixed motives. But one reason they do so is to make people feel safe and respected within online communities. If end-users feel safe, they will continue to participate, post content, and make the platform part of their daily lives. This allows platform owners to be profitable. But it also helps foster a constant, vibrant flow of ideas and opinions.” (BALKIN, Jack M. Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation. **University of California Davis Law Review**. v. 51, n. 3, p. 1149–1210, 2018. Disponível em: https://lawreview.law.ucdavis.edu/issues/51/3/Essays/51-3_Balkin.pdf. Acesso em: 04/11/2021. p. 1195).

²⁷⁵ BALKIN, Jack M. Free Speech is a triangle. **Columbia Law Review**. v. 118, p. 2011–2055, 2018. p. 2021. Disponível em: <https://openyls.law.yale.edu/bitstream/handle/20.500.13051/17953/Balkin%2C%20Free%20Speech%20is%20a%20Triangle.pdf?sequence=1&isAllowed=y>. Acesso em: 04/11/2021.

²⁷⁶ Tradução livre do original: “[...] platforms should be thought of as operating as the New Governors of online speech. These New Governors are part of a new triadic model of speech that sits between the state and speakers/publishers. They are private, self-regulating entities that are economically and normatively motivated to reflect the democratic culture and free speech expectations of their users.” KLONICK, Kate. The New Governors: The People, Rules and Processes Governing Online Speech. **Harvard Law Review**. v. 131, p. 1598–1670. 2018. p. 1603. Disponível em: https://harvardlawreview.org/wp-content/uploads/2018/04/1598-1670_Online.pdf. Acesso em: 28/05/2022.

de mídia social impulsionaram o cumprimento de seus acordos de termos de serviço, além de atender a diversas pressões oriundas de estados-nação para controlar ou refrear discursos *online*, considerados impróprios ou ilegais por esses países.²⁷⁷

Essa construção representa um dos lados da visão triangular da liberdade de expressão de Balkin, a qual não é mais modelo diádico; desta forma, o Estado assume um lado; os falantes e editores, o outro. O novo modelo é pluralista: em uma vertente, o Estado e entidades supranacionais; na outra, o usuário, e, por fim, no segundo lado estão as empresas que operam a infraestrutura digital – como *Facebook*, *YouTube*, *Twitter*, *Google*, seus concorrentes e sucessores – as quais são as novas governantes da expressão digital. Essas organizações desenvolvem e impõem seus termos de serviço, contratos de licença de usuário final e políticas internas da empresa, aplicando-os a usuários finais e membros da comunidade.²⁷⁸

Com efeito, esse predomínio no mercado digital, impulsionado pelo desenvolvimento de tecnologias, algoritmos para processar, lucrar sobre a quantidade de informações e dados gerados, transcendeu o debate de uma questão econômica. Essa união de aparatos tecnológicos e tomadas de decisão por plataformas, alterou, substancialmente, a liberdade econômica para um exercício de poder, até então conferidas às autoridades públicas. Consoante o citado anteriormente, expõe Oreste Policino, “para além dos poderes públicos, a liberdade de conduzir negócios passou agora a uma nova dimensão a do poder privado”.²⁷⁹ Resultando em uma implementação em larga escala de algoritmos pelos atores privados emergentes – os quais passaram a ser imbuídos de poderes não mais os de natureza meramente econômica.

Nossa abordagem é examinar se os Termos de Uso – em especial do *Facebook* – representam ou buscam representar um arranjo de constitucionalização das mídias sociais, e, com isso, observam e se adequam aos ditames dos direitos humanos, especialmente, a

²⁷⁷ KLONICK, Kate. The New Governors: The People, Rules and Processes Governing Online Speech. **Harvard Law Review**. v. 131. p. 1598-1670. 2018. p. 1618-1630. Disponível em: https://harvardlawreview.org/wp-content/uploads/2018/04/1598-1670_Online.pdf. Acesso em: 28/05/2022.

²⁷⁸ BALKIN, Jack M. Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation. **University of California Davis Law Review**. v. 51, n. 3, p. 1149–1210, 2018. p. 1149-1150. Disponível em: https://lawreview.law.ucdavis.edu/issues/51/3/Essays/51-3_Balkin.pdf. Acesso em: 04/11/2021.

²⁷⁹ Tradução livre do original: “Beyond public powers, the freedom of conduct business has now turned into a new dimension, namely that of private power, which [...]”. POLLICINO, Oreste. **Digital Private Powers Exercising Public Functions: The Constitutional Paradox in the Digital Age and its Possible Solutions**. 2021. Disponível em: https://echr.coe.int/Documents/Intervention_20210415_Pollicino_Rule_of_Law_ENG.pdf. Acesso em: 27/10/2021.

liberdade de expressão.²⁸⁰ Nos apoiaremos, tanto no estudo de Edoardo Celeste numa visão de constitucionalização desses termos de uso ou serviço, como no estudo realizado pelo projeto “Termos de Uso e Direitos Humanos”, produzido pelo Centro de Tecnologia e Sociedade da Escola de Direito do Rio de Janeiro da Fundação Getúlio Vargas (CTS/FGV), o qual analisou o impacto dos Termos de Uso no exercício dos direitos humanos à liberdade de expressão, privacidade e processo justo *online*. Sua estruturação contou com a parceria da *Coalizão Dinâmica sobre a Responsabilidade das Plataformas* (DCPR), criada no seio do Fórum de Governança da *Internet* das Nações Unidas.²⁸¹ Por fim, apresentamos as principais atualizações na Política de Privacidade e nos Termos de Serviço do *Facebook*, os quais entraram em vigor em 26 de julho de 2022.

Ademais, serão abordados os principais pontos do primeiro Relatório Anual de Direitos Humanos da *Meta Platforms, Inc.*,²⁸² do qual o *Facebook* integra, lançado em julho de 2022; bem como a recente *Declaração para o Futuro da Internet*, lançada em 28 abril de 2022, propondo um futuro para a *Internet* que seja aberto, livre, global, interoperável, confiável e segura.

2.4.1. Li e Aceito: termos de uso possuem tom constitucional

²⁸⁰ O trabalho situa na análise da atuação da empresa privada de mídia social na governança do discurso *online*. Contudo, é cediço a existência de várias vertentes de estudos sobre a Governança, ou Regulação da *Internet*. Nesse ponto, é salutar a reflexão de Rory Van Loo: “Um debate de décadas nos estudos de direito corporativo e administrativo diz respeito a “como melhor aproveitar os interesses privados dos participantes da empresa para servir ao interesse público”. Historicamente, o ponto de partida era a esperança de que as empresas se autorregulassem – se não por incentivos de mercado, pelo menos para evitar punições legais por irregularidades. Embora os estudiosos reconheçam a heterogeneidade dos aplicadores privados externos, eles não chegaram a examinar a importância emergente de como as grandes empresas são obrigadas a supervisionar terceiros.” Tradução livre do original: “A decades-long debate in both corporate and administrative law scholarship concerns “how best to tap the private interests of enterprise participants to serve the public interest.” Historically, the starting point was the hope that firms would self-regulate—if not for market incentives, then to avoid legal punishment for wrongdoing.³⁵ Although scholars recognize the heterogeneity of external private enforcers,³⁶ they have stopped short of examining the emerging importance of how large firms are required to oversee third parties.” (LOO, Rory Van. **The New Gatekeepers: Private Firms as Public Enforcers**. 106 **Virginia Law Review**, 467, p. 01-49, 2020. Disponível em: https://scholarship.law.bu.edu/faculty_scholarship/800. Acesso em: 03/11/2021).

²⁸¹ VENTURINI, Jamile; LOUZADA, Luiza; MACIEL, Marília. [et. al.]. **Termos de uso e direitos humanos: uma análise dos contratos das plataformas online**. 1 ed. Rio de Janeiro: Revan, 2019. p. 9.

²⁸² Cumpre esclarecer que até aqui o estudo fez referência a rede social Facebook, contudo em 28 de outubro de 2021, o *Facebook, Inc.* mudou seu nome para *Meta Platforms, Inc.* Com isso, passou-se a chamar *Meta*, a empresa que possui em seu catálogo diversas plataformas, como: *Facebook, Messenger, Instagram, WhatsApp, Meta Quest, Workplace, Meta Portal, Metaverso*, dentre outros. E a referência a *Facebook*, quando se refere a plataforma de mídia social, e não à empresa como um todo. Para melhor entendimento conferir: META. **Meta Human Rights Report: Insights and Actions 2020-2021**. Julho de 2022. p. 1-83, Disponível em: https://about.fb.com/wp-content/uploads/2022/07/Meta_Human-Rights-Report-July-2022.pdf. Acesso em: 20/07/2022.

É bastante comum a tentativa de cadastro de qualquer indivíduo em alguma plataforma *online*, e, a partir desse momento, seja submetida a concordar com os termos de uso ou termos de serviço, geralmente proposto com os dizeres: “Li e aceito os termos de uso”. Não há nenhuma espécie de negociação ou abrandamento sugerido ao usuário, sendo assim, ou aceita os termos constituídos e propostos unilateralmente pela empresa ou não aceita, por consequência, não estará habilitado para completar o cadastro e utilizar a referida plataforma, sendo convidado se retirar. Os termos são confeccionados com passe no “*take it or leave it*” – “é pegar ou largar” –, esses Termos de Uso são os contratos de formulário padrão, os quais são os mais utilizados na história mundial.²⁸³ Na verdade, consistem em uma condição de acesso a dados digitais e plataformas baseados em informações.

A ressalva para a utilização desse formato pelas empresas de mídia social, se agrava, ainda mais, em face da aceitação do usuário recair sobre temas sensíveis, como coleta, uso e compartilhamento de dados, incidindo sobre a privacidade, proteção de dados e liberdade de expressão.²⁸⁴ Com isso, extrapola de uma simples relação de consumo, passando a incidir em questões de aplicação de direitos humanos.²⁸⁵

Retoma-se, portanto, a retórica de Zuckerberg, o qual apresenta os termos do *Facebook* como se fosse uma lei fundamental, a qual protege os direitos dos usuários de forma universal. Circunstância abordada por Celeste, afirmando representar mecanismo de

²⁸³ RUSTAD, Michael L.; ONUFRIO, Maria Vittoria. Reconceptualizing consumer terms of use for a globalized knowledge economy. **University of Pennsylvania. Journal of Business Law**. v. 14, n. 4, p. 1085-1190, 2011. p. 1086.

²⁸⁴ Nesse sentido, Gilmar Mendes e Victor Fernandes, tem defendido que aceitar obrigações em matéria de restrições à liberdade de conteúdo, revela contornos mais complexos, afirmando que “o cardápio de obrigações recíprocas é definido não apenas nos ‘Termos de Uso’ ou ‘Termos de Serviços’ da rede social, mas também, de forma complementar, em documentos independentes por vezes intitulados de ‘Declarações de Direitos’, ‘Diretrizes da Comunidade’ ou ‘Regras da Plataforma’. Embora desprovidos de força vinculante, esses documentos representam normas estatutárias acessórias dos contratos privados e funcionam como uma espécie de ‘regramento da comunidade’, cuja interpretação é em geral determinante para as decisões de exclusão de conteúdos e usuários da plataforma.” (MENDES, Gilmar Ferreira; FERNANDES, Victor Oliveira. Eficácia dos direitos fundamentais nas relações privadas da internet: o dilema da moderação de conteúdo em redes sociais na perspectiva comparada Brasil-Alemanha. **Revista de Direito Civil Contemporâneo**. v. 31, ano 9, p. 33-68. São Paulo: Ed. RT, abr./jun. 2022. p. 5 Disponível em: <http://revistadotribunais.com.br/maf/app/document?stid=st-rql&marg=DTR-2022-9857>. Acesso em: 05/08/2022).

²⁸⁵ Como explica, Gunther Teubner: “Sob o pretexto da contratação, os intermediários do mundo cibernético desenvolveram regulamentações privadas autoritárias que não mais podem ser qualificadas como relações contratuais individuais, mas possuem praticamente todas as características da legislação geral.” Tradução livre do original: “Under the guise of contracting, the intermediaries of the cyberworld have developed authoritative private regulations that no longer can be qualified as individual contractual relationships, but have practically all the characteristics of general legislation.” (TEUBNER, Gunther. Horizontal Effects of Constitutional Rights in the Internet: Lergal Case on the Digital Constitution. **The Italian Law Journal**. v. 3, n. 1, p. 193-205, Julho 2017. p. 205. Disponível em: <https://theitalianlawjournal.it/data/uploads/3-italj-1-2017/pdf/teubner.pdf>. Acesso em: 30/05/2022).

constitucionalização no ambiente das mídias sociais,²⁸⁶ aduzindo como outra característica dessa pretensão o modo direcionado ao usuário, empregando o termo “pessoas” para se referir aos usuários.²⁸⁷ Até janeiro de 2022,²⁸⁸ os termos de uso do *Facebook* eram os únicos, dentre as principais redes sociais, denominando-se pelo conceito de “Declaração de Direitos e Responsabilidade”²⁸⁹ – inclusive, constando a seguinte mensagem: “Nossa Declaração de Direitos e Responsabilidades está mudando. Você pode visualizar nossos novos Termos de Serviço aqui”²⁹⁰; contudo, após as alterações, permaneceu sem modificação profunda no seu teor.²⁹¹

Tal qual na versão anterior, os Termos de Uso do *Facebook* possui um caráter genérico, com regras direcionadas às “pessoas”, como se as normativas fossem destinadas às

286 CELESTE, Edoardo. Terms of Service and Bill of Rights: New Mechanisms of Constitutionalisation in the Social Media Environment? **Internacional Review of Law, Computers & Technology**. v. 33, n. 02, p. 122-138. 2019. p. 2. Disponível em: <https://doi.org/10.1080/13600869.2018.1475898>. Acesso em 26/06/2022.

287 Por exemplo: Padrão de Comunidade do Facebook – Discurso de ódio: fundamentos da política: “Acreditamos que as pessoas se comunicam e se conectam mais livremente quando não se sentem atacadas pelo que são. É por isso que não permitimos discursos de ódio no Facebook. Isso cria um ambiente de intimidação e exclusão que, em alguns casos, pode promover a violência no meio físico. [...] Reconhecemos que as pessoas às vezes compartilham conteúdos que incluem discurso de ódio de outra pessoa para condená-la ou aumentar a conscientização. Em outros casos, o discurso que poderia violar nossas normas pode ser usado de forma autorreferencial ou empoderadora. Nossas políticas visam dar espaço para esses tipos de discurso, mas exigimos que as pessoas indiquem claramente sua intenção. Se a intenção não for clara, poderemos remover o conteúdo.” (Facebook. **Padrões da Comunidade do Facebook: discurso de ódio**. 2022. Disponível em: <https://transparency.fb.com/pt-br/policies/community-standards/hate-speech/>. Acesso em: 25/06/2022).

288 Com a criação da *Meta Platforms*, houve alteração nos termos de uso, conforme pode ver na declaração de atualização: “A empresa do Facebook agora é Meta. Atualizamos nossos Termos de Uso, Política de Dados e Política de Cookies para refletir o novo nome em 4 de janeiro de 2022. Embora o nome da nossa empresa tenha mudado, continuamos a oferecer os mesmos produtos, incluindo o aplicativo do Facebook da Meta. Nossa Política de Dados e Termos de Serviço permanecem em vigor, e essa mudança de nome não afeta a forma como usamos ou compartilhamos dados. Saiba mais sobre o Meta e nossa visão para o metaverso.” Tradução livre do original: “The Facebook company is now Meta. We’ve updated our Terms of Use, Data Policy, and Cookies Policy to reflect the new name on January 4, 2022. While our company name has changed, we are continuing to offer the same products, including the Facebook app from Meta. Our Data Policy and Terms of Service remain in effect, and this name change does not affect how we use or share data. Learn more about Meta and our vision for the metaverse.” (FACEBOOK. **Terms of Service**. 2022. Disponível em: https://www.facebook.com/legal/terms/plain_text_terms. Acesso em: 25/06/2022).

289 Importante a previsão no novo Termo de Uso, prevendo a denominação anterior, vejamos: “Estes Termos (anteriormente conhecidos como Declaração de Direitos e Responsabilidades) constituem o acordo integral entre você e a Meta Platforms, Inc. em relação ao uso de nossos Produtos. Eles substituem quaisquer acordos anteriores.” (FACEBOOK. **Terms of Service**. 2022. Disponível em: https://www.facebook.com/legal/terms/plain_text_terms. Acesso em: 25/06/2022).

290 FACEBOOK. **Declaração de Direitos e Responsabilidades**. 2015. Disponível em: <https://pt-br.facebook.com/legal/terms/previous>. Acesso em: 25/06/2022.

291 Essa visão também é compreendida por Nino Shengelia, que afirma: “Os termos do Facebook, em contraste com outras plataformas de mídia social, não parecem ser um simples contrato entre as partes em privado. É possível perceber a disposição da empresa em oferecer aos seus termos um tom mais constitucional.” Tradução livre do original: “Facebook terms, contrasting with other social media platforms, do not seem to be a simple contract between parties in private. It is possible to view willingness of the company to offer its terms a more constitutional tone.” (SHENGELIA, Nino. Constitutionalizing Role of Terms of Service of Social Media Platforms and Proposed Social Media User Bills of Rights. **International Journal of Innovative Technologies in Social Science**. v. 7, n. 28. 2020. p. 4. Disponível em: <https://rsglobal.pl/index.php/ijitss/article/view/1726/1607>. Acesso em 26/06/2021).

comunicações no mundo externo, independentemente do *Facebook*. Outra característica apontada por Celeste, estava representada no nono Princípio do *Facebook*, o qual preceituava a necessidade de quaisquer alterações nos termos de serviço ser submetida à revisão e comentários dos usuários. Com a alteração promovida em 2022, à luz do subitem “1. *Atualizando nossos Termos*”, do item “4. *Disposições adicionais*”, tem-se a empresa justificando justifica, que atualmente só fará alterações em caso de disposições inapropriadas ou incompletas, e somente se essas possíveis alterações forem razoáveis, levando em conta seus interesses ou para fins de segurança, ou para cumprir determinação legal.

Acrescenta, ainda, a promessa de realização de uma notificação prévia a “você” – o usuário – no prazo de 30 dias da realização da alteração, sendo oportunizado a “revisão” antes de entrar em vigor.²⁹² Apesar de não constar a previsão de sujeição a comentários dos usuários, como era previsto no texto anterior, percebe-se a persistência da participação do usuário no processo de alteração dos termos.

Desta forma, reúne características diversas de outras redes sociais, levando a concluir que os termos de uso do *Facebook* transcendem de um mero contrato entre particulares, conforme pontua Celeste:

Em conclusão, ao contrário de outros sites de redes sociais, os termos do Facebook não aparecem como um mero contrato entre particulares; inversamente, é possível perceber a disposição do Facebook em dar um tom constitucional aos seus termos. Isso fica evidente se considerarmos tanto a forma quanto o conteúdo dos termos do

²⁹² Confira o teor integral: “1. Atualizando nossos Termos. Trabalhamos constantemente para melhorar nossos serviços e desenvolver novos recursos para tornar nossos Produtos melhores para você e nossa comunidade. Como resultado, podemos precisar atualizar estes Termos de tempos em tempos para refletir com precisão nossos serviços e práticas, para promover uma experiência segura em nossos Produtos e serviços e/ou para cumprir a lei aplicável. Só faremos alterações se as disposições não forem mais apropriadas ou se estiverem incompletas, e somente se as alterações forem razoáveis e levarem em consideração seus interesses ou se as alterações forem necessárias para fins de segurança ou para cumprir a lei aplicável. Notificaremos você (por exemplo, por e-mail ou por meio de nossos Produtos) pelo menos 30 dias antes de fazermos alterações nestes Termos e lhe daremos a oportunidade de revisá-los antes que entrem em vigor, a menos que as alterações sejam exigidas por lei. Assim que quaisquer Termos atualizados entrarem em vigor, você ficará vinculado a eles se continuar a usar nossos Produtos. Esperamos que você continue usando nossos Produtos, mas se você não concordar com nossos Termos atualizados e não quiser mais fazer parte da comunidade Meta, poderá excluir sua conta a qualquer momento.” Tradução livre do original: “1. Updating our Terms. We work constantly to improve our services and develop new features to make our Products better for you and our community. As a result, we may need to update these Terms from time to time to accurately reflect our services and practices, to promote a safe and secure experience on our Products and services, and/or to comply with applicable law. We will only make any changes if the provisions are no longer appropriate or if they are incomplete, and only if the changes are reasonable and take due account of your interests or if the changes are required for safety and security purposes or to comply with applicable law. We will notify you (for example, by email or through our Products) at least 30 days before we make changes to these Terms and give you an opportunity to review them before they go into effect, unless changes are required by law. Once any updated Terms are in effect, you will be bound by them if you continue to use our Products. We hope that you will continue using our Products, but if you do not agree to our updated Terms and no longer want to be a part of the Meta community, you can delete your account at any time.” (FACEBOOK. **Terms of Service**. 2022. Disponível em: https://www.facebook.com/legal/terms/plain_text_terms. Acesso em: 25/06/2022).

Facebook. Do ponto de vista formal, o Facebook libera seus termos da verborragia característica dos acordos privados para adotar o jargão tradicional do direito constitucional. Do ponto de vista substantivo, o Facebook adapta princípios constitucionais tradicionais, como a liberdade de expressão ou o direito à privacidade e proteção de dados, ao contexto de sua plataforma. Por último, mas não menos importante, sugere uma ideia de “democracia” ao permitir que os usuários participem do processo de definição dos termos.²⁹³

Nessa constatação levantada por Celeste, em reconhecer um tom constitucional nos termos do *Facebook*,²⁹⁴ utiliza o conceito de constitucionalização construído por Anne Peters, definido como “uma abreviação para o surgimento do direito constitucional dentro de uma determinada ordem jurídica.”²⁹⁵ Esse conceito leva em consideração a existência de uma constituição ou lei constitucional a qual existisse em um processo com o passar do tempo. Bem como, um texto legal – ou vários textos – com a possibilidade de características constitucionais num processo cíclico de retroalimentação, fortificando e ganhando contornos constitucionais.²⁹⁶

Peters dividiu em três partes as noções de constituições, quanto às funções e conteúdo, devendo ter um texto, para ser considerada uma lei constitucional ou constituição, a saber: a noção mais ampla da constituição organizando e institucionalizando uma política; a noção mais estrita e funcional da constituição, denotando regras e princípios, cumprindo funções constitucionais típicas; por fim, a mais restrita, denominada *noção legitimista* de uma constituição, respaldada no Art. 16 da Declaração dos Direitos do Homem e do Cidadão de

²⁹³ CELESTE, Edoardo. Terms of Service and Bill of Rights: New Mechanisms of Constitutionalisation in the Social Media Environment? **Internacional Review of Law, Computers & Technology**. v. 33, n. 02, p. 122-138. 2019. p. 3-4. Disponível em: <https://doi.org/10.1080/13600869.2018.1475898>. Acesso em 26/06/2022.

²⁹⁴ No mesmo sentido, tem argumentado Nino Shengelia, afirmando que “os documentos do termo de serviço são documentos constitucionais essenciais sobre como os ambientes sociais compartilhados são governados e constituídos. As documentações dos termos de serviço oferecem aos seus operadores uma quantidade considerável de poder. Esses documentos, para grandes plataformas corporativas, em particular, são escritos de uma maneira que visa salvaguardar os interesses comerciais dos fornecedores de plataforma.” Tradução livre do original: “Service term documents are constitutional documentation that is essential on how shared social environments are governed as well as constituted. Service term documentations offer their operators a considerable amount of power. These documents, for corporate, large platforms, in particular, are written in a manner that aims at safeguarding the platform providers’ commercial interests.” (SHENGELIA, Nino. Constitutionalizing Role of Terms of Service of Social Media Platforms and Proposed Social Media User Bills of Rights. **International Journal of Innovative Technologies in Social Science**. v. 7, n. 28. 2020. p. 6. Disponível em: <https://rsglobal.pl/index.php/ijitss/article/view/1726/1607>. Acesso em 26/06/2021).

²⁹⁵ Tradução livre do original: “‘Constitutionalization’ is shorthand for the emergence of constitutional law within a given legal order.” (PETERS, Anne. Compensatory Constitutionalism: The Function and Potential of Fundamental International Norms and Structures. **Leiden Journal of International Law**. v. 19, n. 3, p. 579-610, 2006. p. 582. Disponível em: https://edoc.unibas.ch/5309/1/20100219153135_4b7ea0c729a55.pdf. ou doi:10.1017/S0922156506003487. Acesso em: 15/05/2022.)

²⁹⁶ PETERS, Anne. Compensatory Constitutionalism: The Function and Potential of Fundamental International Norms and Structures. **Leiden Journal of International Law**. v. 19, n. 3, p. 579-610, 2006. p. 582. Disponível em: https://edoc.unibas.ch/5309/1/20100219153135_4b7ea0c729a55.pdf. ou doi:10.1017/S0922156506003487. Acesso em: 15/05/2022.

1789,²⁹⁷ segundo o qual os direitos humanos e a separação de poderes são os elementos necessários de uma constituição.²⁹⁸

Em resumo, Celeste adotando a noção legitimista, constatou que os termos de uso não são capazes de desempenhar as duas funções constitucionais supremas de proteção dos direitos fundamentais e o equilíbrio dos poderes dos atores envolvidos. Apresenta uma série de questões dificultando a funcionalidade dos termos de serviço como um novo mecanismo de constitucionalização,²⁹⁹ surgindo implicações tanto através do auto constrição, expressão do constitucionalismo digital.³⁰⁰

Contudo, conclui-se, pois, com o surgimento das mídias sociais “subverteu o equilíbrio constitucional, afetando significativamente os direitos fundamentais dos indivíduos e alterando o equilíbrio dos poderes.”³⁰¹ Com isso, os arranjos de constitucionalização, além do Estado ou atividade tipicamente estatal, poderiam ser utilizados para “compensar o fracasso dos instrumentos constitucionais centrados no Estado em enfrentar os desafios jurídicos gerados pelo advento das mídias sociais.”³⁰² Identificou, à luz da teoria do constitucionalismo compensatório, desenvolvida por Peters, a possibilidade da “configuração

²⁹⁷ Art. 16º A sociedade em que não esteja assegurada a garantia dos direitos nem estabelecida a separação dos poderes não tem Constituição.

²⁹⁸ PETERS, Anne. Compensatory Constitutionalism: The Function and Potential of Fundamental International Norms and Structures. **Leiden Journal of International Law**. v. 19, n. 3, p. 579-610, 2006. p. 582. Disponível em: https://edoc.unibas.ch/5309/1/20100219153135_4b7ea0c729a55.pdf. ou doi:10.1017/S0922156506003487. Acesso em: 15/05/2022.

²⁹⁹ Em sentido contrário, Poletto e Moraes: “Resta claro que os Padrões do Facebook possuem um tom constitucional e perseguem os princípios constitucionais democráticos, que decorre da pressão jurídica do movimento de constitucionalização do ambiente digital para a proteção dos direitos dos indivíduos frente às regulações das redes sociais.” (POLETTTO, Álerton; MORAIS, Fausto Santos de. A regulação constitucional do discurso de ódio no Facebook. **Revista Direito, Inovação e Regulações**. Cascavel. Jan. 2022. v. 1, n. 1, p. 36-54, 2022. p. 50)

³⁰⁰ Conferir: GILL, Lex. REDEKER, Dennis; e GASSER, Urs. Towards Digital Constitutionalism? Mapping Attempts to Craft an Internet Bill of Rights (November 9, 2015). **Berkman Center Research Publication**. N. 2015-15. Disponível em: <http://dx.doi.org/10.2139/ssrn.2687120>. Acesso em: 30/01/2021; MENDES, Gilmar Ferreira; FERNANDES, Victor Oliveira. Constitucionalismo digital e jurisdição constitucional: uma agenda de pesquisa para o caso brasileiro. **Revista Brasileira de Direito**. v. 16, n. 1, Passo Fundo, 2020; CELESTE, Edoardo. Digital Constitutionalism: Mapping the Constitutional Response to Digital Technology's Challenges. **HIIG Discussion Paper Series**. N. 2018-02, p. 01-24, 2018. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3219905 ou <http://dx.doi.org/10.2139/ssrn.3219905>. Acesso em: 18/02/2021. SUZOR, Nicolas. Digital Constitutionalism: Using the Rule to Evaluate the Legitimacy of Governance by Platforms. **Social Media + Society**. v. 4, N. 3, July 2018. Disponível em: <https://journals.sagepub.com/doi/pdf/10.1177/2056305118787812>. Acesso em: 18/02/2021.

³⁰¹ Tradução livre do original: “The advent of social media has subverted the constitutional equilibrium by significantly affecting individuals’ fundamental rights and altering the balance of powers.” (CELESTE, Edoardo. Terms of Service and Bill of Rights: New Mechanisms of Constitutionalisation in the Social Media Environment? **Internacional Review of Law, Computers & Technology**. v. 33, n. 02, p. 122-138. 2019. Disponível em: <https://doi.org/10.1080/13600869.2018.1475898>. Acesso em 26/06/2022. p. 12.)

³⁰² Tradução livre do original: “[...] compensate the failure of state-centric constitutional instruments to face the legal challenges generated by the advent of social media.” (CELESTE, Edoardo. Terms of Service and Bill of Rights: New Mechanisms of Constitutionalisation in the Social Media Environment? **Internacional Review of Law, Computers & Technology**. v. 33, n. 02, p. 122-138. 2019. p. 12. Disponível em: <https://doi.org/10.1080/13600869.2018.1475898>. Acesso em 26/06/2022).

dos termos de serviço das mídias sociais como mecanismos de constitucionalização no ambiente das mídias sociais podem ser encaradas como uma expressão dessa tendência compensatória.”³⁰³ Assim a utilização de um tom constitucional poderia ser “considerada sintomática da disposição desses documentos em desempenhar tal função compensatória, e a dimensão limitada desse fenômeno como evidência do estado embrionário desse processo.”³⁰⁴

2.4.2. Termos de uso e a liberdade de expressão

Como visto, os *gatekeepers* de autoridade³⁰⁵ – dentre eles o *Facebook* – tem desenvolvido arranjos cujo impacto ocorre, dentre outros aspectos, contra os direitos de liberdade de expressão, informação e opinião. Desencadeando preocupação, de intermediários usados para implementar políticas públicas com supervisão limitada e responsabilização com graves implicações sobre os direitos humanos.³⁰⁶

Tais direitos humanos são centrados no Estado, como principal detentor do dever de proteção, cabendo a este a garantia de os atores privados não violarem os direitos humanos. A relação natural de tais direitos é focada num *efeito vertical* – ou seja, obrigações do Estado para o indivíduo. Mas reconhece um *efeito horizontal*, podendo surgir numa relação entre partes privadas, devendo proteger os direitos humanos no âmbito privado. Isso tem sido

³⁰³ Tradução livre do original: “[...] configuration of social media’s terms of service and the bills of rights of social media users as mechanisms of constitutionalisation in the social media environment could be envisaged as an expression of this compensatory trend.” (CELESTE, Edoardo. Terms of Service and Bill of Rights: New Mechanisms of Constitutionalisation in the Social Media Environment? **Internacional Review of Law, Computers & Technology**. v. 33, n. 02, p. 122-138. 2019. p. 12. Disponível em: <https://doi.org/10.1080/13600869.2018.1475898>. Acesso em 26/06/2022).

³⁰⁴ Tradução livre do original: “[...] could be regarded as symptomatic of the willingness of these documents to play such a compensatory function, and the limited dimension of this phenomenon as an evidence of the embryotic state of this process.” (CELESTE, Edoardo. Terms of Service and Bill of Rights: New Mechanisms of Constitutionalisation in the Social Media Environment? **Internacional Review of Law, Computers & Technology**. v. 33, n. 02, p. 122-138. 2019. p. 12. Disponível em: <https://doi.org/10.1080/13600869.2018.1475898>. Acesso em 26/06/2022).

³⁰⁵ Classificação proposta por Emily Laidlaw, das empresas de Internet no controle da esfera públicas online. De acordo com o impacto democrático, aponta três tipos de *gatekeepers*: os macros gatekeepers que mantêm controle significativo devido os usuários deverem passar por elas para usar a internet. Por exemplo: provedores de serviços de internet, provedores de rede móvel e os principais mecanismos de pesquisa; os gatekeepers de autoridade, que controlam grandes quantidades de tráfego e fluxo de informações, apesar de não serem indispensáveis para acessar a Internet. Por exemplo: Wikipedia e Facebook; por fim, os micros *gatekeepers* que possuem participação importante, facilitando informações e debates democráticos, a exemplo de sites de notícias. (LAIDLAW, Emily B.A framework for identifying Internet information gatekeepers. **International Review of Law, Computers & Technology**. v. 24, n. 3, p. 263-276, 2010. p.271-274. Disponível em: https://ueaeprints.uea.ac.uk/id/eprint/29193/1/A_Framework_for_Identifying_Internet_Information_Gatekeepers.pdf. Acesso em: 30/06/2022).

³⁰⁶ A propósito conferir: KUCZERAWY, Aleksandra. Responsabilidade intermediária & liberdade de expressão: Desenvolvimentos recentes na iniciativa de aviso & ação da UE. **Computer Law & Security Review: The International Journal of Technology**. v. 31, n. 1, p. 46-56, 2015. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0267364914001836>. Acesso em: 30/06/2022.

objeto de estudos relacionados as liberdades *online*, com novos meios de interferência estatal nos direitos humanos, como novos meios de restrição de conteúdo, vigilância ou envolvimento de empresas de *Internet* na aplicação da lei.³⁰⁷

No ambiente *online*, verifica-se que a interferência na liberdade de expressão pode ser exercida de formas diversas, como a fragilidade e a inadequação do direito à privacidade, da proteção de dados, repercutindo a impossibilidade de disseminação de opiniões e informações, podendo ainda implicar medidas que removam o conteúdo ou desabilitem a capacidade dos usuários, com exclusão de contas ou bloqueio.³⁰⁸

Nesse contexto, percebe-se que implicações com os direitos humanos, especialmente a liberdade de expressão, são, frequentemente, encontradas no ambiente *online* na primazia do controle pertence a atores privados. Assim, são necessários novos arranjos para garantir a higidez do discurso *online*, Balkin então destaca como essa infraestrutura digital e os serviços *online* estão sob controle de empresas privada. A moderação (remoção ou bloqueio de acesso ao conteúdo) desse debate é necessária à cooperação entre esses prestadores de serviços ou coerção exercida sobre eles através de técnicas, denominada de "nova escola".³⁰⁹ E aponta três características dessa escola: (i) censura colateral, em que o Estado regula uma das partes – o intermediário – para controlar outra, o orador, (ii) cooperação pública-privada e cooptação, como o governo não possui a infraestrutura da liberdade de expressão, ele precisa coagir ou cooptar proprietários privados para auxiliar na regulação e vigilância do discurso e (iii) novas formas de restrições prévias digitais.³¹⁰

Busca-se, nesse ponto, a análise da violação dos direitos humanos no plano *horizontal* – entre intermediários e usuário. Pois, a potencial violação da liberdade de expressão e informação é atribuível ao Estado e o conflito de direitos humanos, portanto, um conflito

³⁰⁷ JØRGENSEN, Rikk Frank. Human Rights and Private Actors in the Online Domain. In LAND, Molly K.; ARONSON, Jay D; (Eds.), **New Technologies for Human Rights Law and Practice**. p. 243-269. Cambridge: Cambridge University Press. 2018. p. 253. Disponível em: doi:10.1017/9781316838952.011 ou https://www.cambridge.org/core/services/aop-cambridge-core/content/view/AEB97862FED32E804B078A41AFAC97E8/9781107179639c11_243-269.pdf/human-rights-and-private-actors-in-the-online-domain.pdf. Acesso em: 29/10/2021.

³⁰⁸ JØRGENSEN, Rikk Frank; PEDERSEN, Anja Møller. Online Service Providers as Human Rights Arbiters. In: TADDEO, Mariarosalia, FLORIDI, Luciano. [Eds]. **The Responsibilities of Online Service Providers**. **Law, Governance and Technology Series**, v. 31. p. 179-199. Springer, Cham. 2017, Disponível em: https://doi.org/10.1007/978-3-319-47852-4_10. Acesso em: 31/10/2021.

³⁰⁹ BALKIN, Jack M. Old-School/New-School Speech Regulation. **Harvard Law Review**. v. 127, n. 8, p. 2296-2342., 2014. p. 2298. Disponível em: https://openyls.law.yale.edu/bitstream/handle/20.500.13051/4406/127_Harv_L_Rev._2296_2014.pdf?sequence=2&isAllowed=y. Acesso em: 04/11/2021.

³¹⁰ BALKIN, Jack M. Old-School/New-School Speech Regulation. **Harvard Law Review**. v. 127, n. 8, p. 2296-2342. 2014. p. 2299. Disponível em: https://openyls.law.yale.edu/bitstream/handle/20.500.13051/4406/127_Harv_L_Rev._2296_2014.pdf?sequence=2&isAllowed=y. Acesso em: 04/11/2021.

vertical entre o Estado e o intermediário, ou ainda, com o orador ou os usuários finais. Nesses casos, quando a interferência é baseada na lei e/ou incentivada pelo próprio Estado – correção – a potencial violação da liberdade de expressão e informação é atribuível ao Estado.³¹¹

Quando há conflito no plano *horizontal*, no qual reside a preocupação, pois a regulações voluntária – autorregulação – incide nesse plano, entre o intermediário que impõe a medida e o usuário – quem está sendo bloqueado ou negado o acesso. Jørgensen e Pedersen pontuam que os conflitos horizontais devem ser resolvidos por meio de obrigações positivas oriundas dos direitos humanos do Estado. Logo, o que determina o resultado é se o “Estado estabeleceu um equilíbrio justo entre a liberdade do intermediário de conduzir um negócio (prestação de serviços de *internet*) e o direito à liberdade de expressão e informação do orador ou usuário final”.³¹²

O foco atual, que está no centro do debate, são as responsabilidades de direitos humanos das empresas de *internet*, que tem pressionado esses atores privados, tanto para implementar políticas públicas (combater conteúdo ilegal), quanto respeitar os direitos humanos. Percebe-se uma movimentação positiva nesse sentido, por exemplo a edição do primeiro relatório anual sobre direitos humanos pelo *Meta*, detentora do *Facebook*.

Nessa perspectiva, apontamos como meio alternativo, para alcançar a responsabilidade dos direitos humanos pelos atores sociais, a observância dos “Princípios Orientadores sobre Negócios e Direitos Humanos das Nações Unidas” (Conselho de Direitos Humanos das Nações Unidas, 2011). É constituído por três pilares mestres: 'Proteger', 'Respeito' e 'Remédio'.

O *primeiro pilar*, descreve o dever do Estado de proteger contra os abusos de direitos humanos por terceiros, incluindo empresas privadas, que exige que os Estados tomem medidas adequadas para prevenir, investigar, punir e corrigir violações de direitos humanos cometidas por atores privados através de políticas, legislação e regulamentos eficazes e

³¹¹ Cumpra deixar assente, que não se ocupa nesse estudo o exame direto dos importantes tipos e formas de regulação, o que demandaria um estudo próprio, dado a extensão e complexidade sobre o tema, conferir os estudos: RIBEIRO, Keila Cristina De Lima Alencar. **Liberdade de expressão e fake news: uma análise acerca da possibilidade de regulação das redes sociais à luz do entendimento do Supremo Tribunal Federal**. 2021. 174 f. Dissertação (Mestrado em Direito Constitucional) – Instituto Brasileiro de Ensino, Desenvolvimento e Pesquisa, Brasília, 2021. PINHO FILHO, José Célio Belém de. **Desinformação e regulação de redes sociais digitais**. 2021. 170 f. Dissertação (Mestrado Profissional em Direito, Justiça e Desenvolvimento). Instituto Brasileiro de Ensino, Desenvolvimento e Pesquisa, Brasília, 2021.

³¹² JØRGENSEN, Rikk Frank; PEDERSEN, Anja Møller. Online Service Providers as Human Rights Arbiters. In: TADDEO, Mariarosaia, FLORIDI, Luciano. [Eds]. *The Responsibilities of Online Service Providers*. **Law, Governance and Technology Series**, v. 31. p. 179-199. Springer, Cham. 2017, Disponível em: https://doi.org/10.1007/978-3-319-47852-4_10. Acesso em: 31/10/2021.

julgamentos.³¹³ O *segundo pilar* é a responsabilidade corporativa de respeitar os direitos humanos, o que implica que as empresas privadas devem publicar um compromisso político de respeitar os direitos humanos e agir com a devida diligência, a fim de evitar infringir os direitos humanos de terceiros³¹⁴. O *terceiro pilar* aborda que os Estados devem adotar medidas – judicial ou extrajudicial – proporcionando acesso desses remédios as vítimas de abusos dos direitos humanos.³¹⁵

Em suma, os Estados são obrigados a prevenir violações de direitos humanos por atores privados, e os atores privados têm o dever moral – não legal – de respeitar os direitos humanos. Assim, os Estados devem garantir práticas comerciais compatíveis com os direitos humanos por meio de regulamentação adequada, e cada empresa tem a responsabilidade de acessar seu real impacto nos direitos humanos.³¹⁶

Recentemente, em junho de 2021, completou-se a marca de 10 anos dos Princípios Orientadores da ONU sobre empresas e Direitos Humanos, sendo elaborado um balanço desse período. Constatando, a persistência de abusos relacionados às empresas, devendo assumir atenção prioritária urgente por parte dos Estados e das empresas, verificou-se que, na última década, ressaltou o ponto levantado nos Princípios Orientadores: *abordagens voluntárias por si só não são suficientes*. E que é necessário o aumento das medidas obrigatórias, para acelerar a aceitação e o progresso, bem como, as medidas legais sendo essenciais, mas não suficientes para garantir o respeito dos direitos humanos pelas empresas.³¹⁷

O balanço, realiza um prospecto para os Princípios Orientadores e a diligência em direitos humanos, centrada, primeiramente, naqueles em maior risco, fornecendo um plano para Estados e empresas para recuperação responsável da crise do COVID-19, pois, a

³¹³ CONSELHO DE DIREITOS HUMANOS DAS NAÇÕES UNIDAS. **Princípios Orientadores sobre Negócios e Direitos Humanos das Nações Unidas**. 2011. Disponível em: https://www.ohchr.org/sites/default/files/documents/publications/guidingprinciplesbusinesshr_en.pdf. Acesso em: 20/01/2022. p. 3-12.

³¹⁴ CONSELHO DE DIREITOS HUMANOS DAS NAÇÕES UNIDAS. **Princípios Orientadores sobre Negócios e Direitos Humanos das Nações Unidas**. 2011. Disponível em: https://www.ohchr.org/sites/default/files/documents/publications/guidingprinciplesbusinesshr_en.pdf. Acesso em: 20/01/2022. p. 13-26.

³¹⁵ CONSELHO DE DIREITOS HUMANOS DAS NAÇÕES UNIDAS. **Princípios Orientadores sobre Negócios e Direitos Humanos das Nações Unidas**. 2011. Disponível em: https://www.ohchr.org/sites/default/files/documents/publications/guidingprinciplesbusinesshr_en.pdf. Acesso em: 20/01/2022. p. 27-35.

³¹⁶ JØRGENSEN, Rikk Frank; PEDERSEN, Anja Møller. Online Service Providers as Human Rights Arbiters. In: TADDEO, Mariarosalia, FLORIDI, Luciano. [Eds]. *The Responsibilities of Online Service Providers*. **Law, Governance and Technology Series**, v. 31. p. 179-199. Springer, Cham. 2017, Disponível em: https://doi.org/10.1007/978-3-319-47852-4_10. Acesso em: 31/10/2021.

³¹⁷ CONSELHO DE DIREITOS HUMANOS DAS NAÇÕES UNIDAS. **Guiding Principles On Business And Human Rights At 10: Taking stock of the first decade**. A/HRC/47/39. 2021. Disponível em: <https://www.ohchr.org/sites/default/files/Documents/Issues/Business/UNGPs10/Stocktaking-reader-friendly.pdf>. Acesso em: 10/02/2022. p.25.

pandemia, apesar de seus efeitos funestos, ofereceu oportunidade única para reconstruir melhor.³¹⁸

Ao finalizar, destaca outros grandes desafios globais os quais exigirão de todas as partes interessadas uma abordagem sistêmica, bem como esforços persistentes para alavancar os múltiplos atores além dos Estados que fomentam políticas, práticas e, de fato, regulamentos para moldar os comportamentos de negócios em uma combinação inteligente de medidas. Deve estar no centro dos desenvolvimentos regulatórios, portanto, o respeito empresarial aos direitos humanos, garantindo recuperação inclusiva e a construção de uma sociedade mais resiliente.³¹⁹

Noutra perspectiva, mas ainda no ímpeto de buscar a observância dos direitos humanos por atores globais, pontua-se o resultado da pesquisa realizada pelo CTS/FGV – Centro de Tecnologia e Sociedade da Escola de Direito do Rio de Janeiro da Fundação Getúlio Vargas em parceria com a *Coalizão Dinâmica sobre a Responsabilidade das Plataformas* (DCPR) no qual foi realizada uma análise da adequação dos contratos das plataformas *online* relacionando-os aos direitos humanos, cujos critérios adotados basearam-se na estrutura do “Guia de Direitos Humanos para os Usuários da *Internet*”³²⁰, utilizando-se, também, de princípios internacionais presentes nas principais convenções de proteção dos direitos humanos.³²¹ A análise, por conseguinte, concentrou-se nos direitos à liberdade de expressão, à privacidade e ao processo justo *online*.³²²

Durante a pesquisa foram analisados os Termos de Uso de 50 plataformas. O primeiro critério observado no tópico sobre Liberdade de Expressão foi relativo à *Análise, bloqueio, filtragem ou remoção de conteúdo*. Verificou-se, em 46% das plataformas, a existência de cláusulas sobre alguma espécie de monitoramento, mesmo sem especificá-lo ou fazendo de

³¹⁸ CONSELHO DE DIREITOS HUMANOS DAS NAÇÕES UNIDAS. **Guiding Principles On Business And Human Rights At 10: Taking stock of the first decade.** A/HRC/47/39. 2021. Disponível em: <https://www.ohchr.org/sites/default/files/Documents/Issues/Business/UNGPs10/Stocktaking-reader-friendly.pdf>. Acesso em: 10/02/2022. p.25.

³¹⁹ CONSELHO DE DIREITOS HUMANOS DAS NAÇÕES UNIDAS. **Guiding Principles On Business And Human Rights At 10: Taking stock of the first decade.** A/HRC/47/39. 2021. Disponível em: <https://www.ohchr.org/sites/default/files/Documents/Issues/Business/UNGPs10/Stocktaking-reader-friendly.pdf>. Acesso em: 10/02/2022. p.25.

³²⁰ Consta, que também foram adotados como critérios de análise: “as declarações da Relatoria Especial para a Liberdade de Expressão da Comissão Interamericana de Direitos Humanos da Organização dos Estados Americanos (OEA) e da Relatoria Especial da ONU sobre Liberdade de Opinião e Expressão que tratam especificamente da aplicação dos direitos humanos ao ambiente online.” VENTURINI, Jamile; LOUZADA, Luiza; MACIEL, Marília. [et. al.]. **Termos de uso e direitos humanos: uma análise dos contratos das plataformas online.** 1 ed. Rio de Janeiro: Revan, 2019. p. 24.

³²¹ VENTURINI, Jamile; LOUZADA, Luiza; MACIEL, Marília. [et. al.]. **Termos de uso e direitos humanos: uma análise dos contratos das plataformas online.** 1 ed. Rio de Janeiro: Revan, 2019. p. 23.

³²² Seguindo o objeto desse estudo, apresentaremos somente os resultados quanto aos direitos à liberdade de expressão.

forma vaga. Inobstante, 40% – representada por 20 empresas – apresentaram resultados contraditórios; por outro lado, 8% explicitaram não exercer nenhum tipo de vigilância de conteúdo, excetuando-se os materiais removidos pela violação da sua política.³²³

À frente, em segundo momento, o aspecto verificado foi a respeito de *Informações sobre a denúncia de conteúdos considerados impróprios ou abusivos*; nesse aspecto, 70% das plataformas pesquisadas oferecem informações claras e transparentes sobre o procedimento a ser seguido pelo usuário para contatar a plataforma ao se depararem com conteúdo impróprio ou que esteja violando as suas políticas. Apenas uma apresenta cláusulas contraditórias e 14 plataformas (28%) não possuem cláusulas relevantes.³²⁴

O terceiro faz uma imersão sobre a *Notificação e direito a ser ouvido em caso de remoção de conteúdos*, o qual evidenciou o número ínfimo de 4% das plataformas cujas normas possibilitem o autor apresentar defesa às acusações sobre algum de seus conteúdos ou remoções indevidas. Somente 4 plataformas (8%) apresentaram cláusulas contraditórias e 36% não são totalmente omissas. Insta pontuar o cômputo de 52% das plataformas, as quais afirmam explicitamente que “pode retirar conteúdos devido a notificação de terceiros sem oferecer qualquer justificativa, notificação ou oportunidade de ser ouvido ao usuário que os compartilhou.”³²⁵

Outro critério ponderado foi sobre a *Notificação e possibilidade de questionamento em caso de terminação de contas individuais*; nesse caso, nenhuma afirmou realizá-la, por outro lado, 88% reiteraram o seu poder em excluir a conta de um usuário sem notificação prévia. Além disso, 8% apresentam dados contraditórios e em duas (4%) não constam cláusulas a respeito.³²⁶

Por fim, examinou-se a *Possibilidade de anonimato e do uso de pseudônimos*, na qual 32% das plataformas não autorizam o anonimato ou uso de pseudônimos; entretanto, esse procedimento é realizado por 28% dessas redes. Apenas 8% das plataformas informaram cláusulas contraditórias e 32% não possuíam cláusulas relevantes para esse critério.³²⁷

³²³ VENTURINI, Jamile; LOUZADA, Luiza; MACIEL, Marília. [et. al.]. **Termos de uso e direitos humanos: uma análise dos contratos das plataformas online**. 1 ed. Rio de Janeiro: Revan, 2019. p. 51-53.

³²⁴ VENTURINI, Jamile; LOUZADA, Luiza; MACIEL, Marília. [et. al.]. **Termos de uso e direitos humanos: uma análise dos contratos das plataformas online**. 1 ed. Rio de Janeiro: Revan, 2019. p. 55.

³²⁵ VENTURINI, Jamile; LOUZADA, Luiza; MACIEL, Marília. [et. al.]. **Termos de uso e direitos humanos: uma análise dos contratos das plataformas online**. 1 ed. Rio de Janeiro: Revan, 2019. p. 56-57.

³²⁶ VENTURINI, Jamile; LOUZADA, Luiza; MACIEL, Marília. [et. al.]. **Termos de uso e direitos humanos: uma análise dos contratos das plataformas online**. 1 ed. Rio de Janeiro: Revan, 2019. p. 58.

³²⁷ VENTURINI, Jamile; LOUZADA, Luiza; MACIEL, Marília. [et. al.]. **Termos de uso e direitos humanos: uma análise dos contratos das plataformas online**. 1 ed. Rio de Janeiro: Revan, 2019. p. 59.

Infere-se dos dados demonstrados o oferecimento de garantias à preservação da liberdade de expressão em seus termos de serviços, as quais, em sua maioria, são feitas de forma bastante incipiente e insatisfatória, bem como não observa a necessidade de uma linguagem clara e precisa quanto aos conteúdos sensíveis, a exemplo da possibilidade de justificar, notificar e do direito a ser ouvido em caso de remoções promovidas por iniciativa da própria plataforma, mesmo após notificação de terceiros.

Percebe-se, de mesmo modo, a necessidade de examinar os arranjos das plataformas quanto à caracterização de conteúdos abusivos. Isso porque, como é um instrumento útil e importante na preservação de direitos fundamentais no ambiente *online*, nem sempre é utilizada com efetividade para resguardar a liberdade de expressão, e, com isso, gera um efeito reverso, tornando-se uma brecha permissiva às instituições privadas para a remoção indiscriminada e discricionária de conteúdo.

Outro aspecto relevante é a tendência das plataformas manejarem seus mecanismos de remoção fundamentadas por requisitos mínimos previstos na lei de direitos autorais dos EUA³²⁸; a intenção é evitar responsabilidade indireta por violação a este conteúdo, evidenciando, ainda, a ausência de arranjos internos, os quais possibilitariam recurso em caso de remoções, em razão dos contra-avisos não terem empecilhos para ser arquivados, conforme os termos da *Digital Millenium Copyright Act* - DMCA, independentemente de um reconhecimento explícito dentro dos termos de serviço.³²⁹

Desta forma, quanto ao procedimento de indisponibilizar espaço às postagens ofensivas aos termos estabelecidos pelas plataformas, a notificação e remoção de conteúdo criado por terceiros, em ambiente online, precisa ser analisada à luz dos preceitos sobre a liberdade de expressão, fato dissertado por Kanayama ao realizar uma abordagem ampla e profunda sob a ótica do Marco Civil da *Internet*, relata-se:

[...] o procedimento de notificação e retirada, como o conhecemos hoje – a partir de sua estruturação nos EUA, sua aplicabilidade nos Tribunais brasileiros e, por fim, sua previsão no Anteprojeto de Reforma à LDA – não satisfaz a condicionante de respeito à liberdade de expressão prevista no § 2º, art. 19, do Marco Civil da Internet. Condicionante que, conforme defendemos ao longo deste trabalho, considerando a natureza de Estatuto do MCI no ordenamento – verdadeira bússola para as demais leis que lidarem com questões relativas ao ambiente digital – e a

³²⁸ Apesar dessa prática, cumpre observar que no Brasil, o Marco Civil da Internet – Lei 12.965/2014 – prevê no Art. 20, o dever dos provedores de aplicações de Internet de notificar os usuários cujo conteúdo for tornado indisponível e de fornecer informações que permitam os usuários exercerem os seus direitos de contestação e apresentação de uma defesa.

³²⁹ VENTURINI, Jamile; LOUZADA, Luiza; MACIEL, Marília. [et. al.]. **Termos de uso e direitos humanos: uma análise dos contratos das plataformas online**. 1 ed. Rio de Janeiro: Revan, 2019. p. 99.

necessidade do diálogo das fontes, não deve ser mera opção no trato legislativo das "infrações" aos direitos autorais na Internet.³³⁰

Além da problemática de proporcionar um ambiente sem violações ao direito à liberdade de expressão, bem como outros direitos fundamentais no ambiente *online*, é imprescindível ressaltar, mesmo com questões a serem melhor trabalhadas, a importância do *acesso ao conteúdo da Internet* tem repercutido nas organizações internacionais em busca de garantir a acessibilidade a nível universal³³¹. Por esse prisma, em um dos seus relatórios, o Relator Especial La Rue salientou ser prioridade dos Estados assegurar a circulação das informações contidas no ambiente virtual à totalidade de seu território³³². Sob a mesma visão, quatro relatores apresentaram uma declaração conjunta em 2011, a qual afirmava: “a efetivação do direito à liberdade de expressão impõe aos Estados a obrigação de promover o acesso universal à *internet*.”³³³ Indubitável evidenciar a *internet* como principal agente à introdução de novos vetores que ameaçam os direitos humanos; entretanto, por outro viés, é um instrumento para amplificar o potencial das pessoas para reivindicar seus direitos, cujo resultado é a facilitação à segurança humana.³³⁴

Abordaremos, como último ponto deste subitem, as recentes alterações realizadas pela *Meta* nas plataformas, sobretudo no *Facebook*, como forma de demonstrar, na prática, os atores sociais na atuação do aperfeiçoamento dos Termos de Uso e Serviço, bem como das

³³⁰ KANAYAMA, Ricardo Alberto. A liberdade de expressão do Marco Civil da Internet e o procedimento de notificação e retirada para as "infrações" aos direitos autorais. **Civilistica.com**. Rio de Janeiro, a. 10, n. 1, 2021. Disponível em: <http://civilistica.com/a-liberdade-de-expressao-do-marco-civil/>. Acesso em: 19/11/2021.

³³¹ Sobre o direito fundamental de acesso à internet, conferir: ARROYO, César Landa. Derecho fundamental al Internet: contenido esencial. **Pensar**. Fortaleza. v. 23, n. 4, p. 1-22, 2018. Disponível em: <https://periodicos.unifor.br/rpen/article/view/8447/pdf>. Acesso em: 24/10/2021; BACA, Vilmar Villena Fernández. Derecho fundamental de acceso a internet: Um acercamiento preliminar a sus contenidos. **Gaceta Constitucional**. N. 163, p. 194-204, 2021. Disponível em: https://www.academia.edu/50343496/Derecho_fundamental_de_acceso_a_internet_Un_acercamiento_preliminar_a_sus_contenidos. Acesso em 26/08/2021.

³³² Consta no item Recomendação, parágrafo 87, que: “O Relator Especial enfatiza que o acesso à informação, a capacidade de exercer o direito à liberdade de expressão e a participação que a Internet proporciona a todos os setores da sociedade é essencial para uma sociedade verdadeiramente democrática.” Tradução livre do original: “The Special Rapporteur emphasizes that the access to information, the ability to exercise the right to freedom of expression and the participation that the Internet provides to all sectors of society is essential for a truly democratic society.” (LA RUE, Frank. **Promotion and protection of the right to freedom of opinion and expression**. 2011. Disponível em: <https://www.ohchr.org/sites/default/files/Documents/Issues/Opinion/A.66.290.pdf>. Acesso em: 20/06/2022, p. 22)

³³³ LIGABO, Ambeyi; DUVE, Freimut; BERTONI, Eduardo. **International mechanisms for promoting freedom of expression**. 2011. Disponível em: www.osce.org/files/f/documents/4/0/28235.pdf. Acesso em 20/06/2022.

³³⁴ KETTEMANN, Matthias C. **The Normative Order of the Internet: A Theory of Rule and Regulation** Online. Oxford: Oxford University Press. 2020. Disponível em: <https://nbn-resolving.org/urn:nbn:de:0168-ss0ar-69191-7>. Acesso em: 26/06/2021. p. 36.

políticas de privacidade, as quais objetivam respaldar os direitos fundamentais, dentre eles a liberdade de expressão.

No dia 26 de julho de 2022, começou a vigorar uma série de alterações na Política de Privacidade e nos Termos de Serviços da *Meta* – mencionando-se o *Facebook*; porém, anteriormente, em 26 de maio de 2022, a *Meta* notificou seus usuários a fim de dar ciência da atualização da nova “Política de Privacidade”³³⁵, a qual era nominada como Política de Dados; a informação abarcava as modificações, as quais foram inspiradas pelas contribuições de “especialistas em privacidade e dos usuários das nossas tecnologias, reescrevemos nossa Política de Privacidade para facilitar a compreensão e refletir os produtos mais recentes que lançamos”.³³⁶ Primando pela abordagem com transparência, chama atenção a algumas partes com redação diferente, esclarecendo que “a *Meta* não coleta, usa ou compartilha seus dados de novas maneiras com base nesta atualização da política e não vende suas informações.”³³⁷ Em síntese, a plataforma informa as principais alterações em três tópicos:

- Reescrevemos e remodelamos nossa Política de Privacidade para facilitar a compreensão e a clareza sobre como usamos suas informações. Nossos Termos de Serviço também estão sendo atualizados, com a finalidade de explicar melhor o que se espera de nós e daqueles que utilizam as nossas plataformas.
- Embora o texto esteja diferente, estas atualizações da Política e dos Termos não permitem à *Meta* coletar, usar ou compartilhar seus dados de novas maneiras.
- Acrescentamos explicações mais detalhadas em nossa Política de Privacidade, inclusive sobre como usamos e compartilhamos as informações com terceiros. Além disso, combinamos a Política com a Central de Privacidade e novos controles para gerenciar sua experiência, por exemplo, quem vê suas publicações e os tópicos sobre os quais você quer ver anúncios.³³⁸

Consoante o supracitado, o objetivo dessa atualização é tornar clara as práticas de dados aplicadas pela plataforma, afirmando a busca da comunicação descomplicada sobre as práticas de dados e as decisões tomadas por usuários na plataforma. Exemplificando-se, consta no Termo de Privacidade a coleta às informações, sejam elas: atividades e informações fornecidas pelo indivíduo; informações sobre amigos, seguidores; sobre contatos; atividades de outras pessoas referentes ao usuário; além de informações de aplicativos, navegadores e dispositivos. Abaixo, elucida-se a previsão no Termo de Privacidade:

³³⁵ META. **Política de Privacidade**. 2022. Disponível em: <https://www.facebook.com/privacy/policy>. Acesso em 20/06/2022.

³³⁶ PROTTI, Michel. O que você precisa saber sobre a atualização da nossa política de privacidade e em nossos termos de serviço. **Meta**. 2022. Disponível em: <https://about.fb.com/br/news/2022/05/o-que-voce-precisa-saber-sobre-a-atualizacao-da-nossa-politica-de-privacidade-e-em-nossos-termos-de-servico/>. Acesso em: 20/06/2022.

³³⁷ PROTTI, Michel. O que você precisa saber sobre a atualização da nossa política de privacidade e em nossos termos de serviço. **Meta**. 2022. Disponível em: <https://about.fb.com/br/news/2022/05/o-que-voce-precisa-saber-sobre-a-atualizacao-da-nossa-politica-de-privacidade-e-em-nossos-termos-de-servico/>. Acesso em: 20/06/2022.

³³⁸ PROTTI, Michel. O que você precisa saber sobre a atualização da nossa política de privacidade e em nossos termos de serviço. **Meta**. 2022. Disponível em: <https://about.fb.com/br/news/2022/05/o-que-voce-precisa-saber-sobre-a-atualizacao-da-nossa-politica-de-privacidade-e-em-nossos-termos-de-servico/>. Acesso em: 20/06/2022.

Quais informações coletamos?

As informações que coletamos e tratamos sobre você dependem do seu uso dos nossos Produtos. Por exemplo, coletamos informações diferentes se você vende móveis no Marketplace ou publica um vídeo do Reels no Instagram. Quando você usa nossos Produtos, coletamos algumas informações sobre você, mesmo que não tenha uma conta.

Estas são as informações que coletamos:

Sua atividade e as informações que você fornece:

Nos nossos Produtos, você pode enviar mensagens, tirar fotos e gravar vídeos, comprar ou vender itens e muito mais. Chamamos de “atividade” tudo o que você pode fazer nos nossos Produtos. Coletamos sua atividade nos nossos Produtos e as informações que você nos fornece, como as seguintes:

- O conteúdo criado, como publicações, comentários ou áudios
- O conteúdo fornecido por meio do nosso recurso de câmera, das configurações do rolo da câmera ou dos nossos recursos habilitados para voz. Saiba mais sobre o que coletamos desses recursos e como usamos as informações da câmera em máscaras, filtros, avatares e efeitos.
- As mensagens enviadas e recebidas, incluindo o conteúdo, sujeitas às leis aplicáveis. Não podemos ver o conteúdo de mensagens criptografadas de ponta a ponta, a menos que os usuários as denunciem para análise. Saiba mais.
- Os metadados sobre conteúdo e mensagens, sujeitos às leis aplicáveis.
- Os tipos de conteúdo visto ou com o qual você interage e o modo como faz isso
- Os aplicativos e recursos usados e quais ações você realiza neles. Veja exemplos.
- As compras ou outras transações efetuadas, incluindo informações de cartão de crédito. Saiba mais.
- As hashtags usadas
- O horário, a frequência e a duração das suas atividades nos nossos Produtos

Os amigos, os seguidores e outras conexões

As informações coletadas sobre amigos, seguidores e outras conexões

Coletamos informações sobre amigos, seguidores, grupos, contas, Páginas do Facebook e outros usuários e comunidades com quem você tem conexão ou interage. Isso inclui a forma como você interage com eles nos nossos Produtos e com quais você interage mais.

As informações que coletamos sobre contatos

Também coletamos as informações dos seus contatos, incluindo nome e endereço de email ou telefone, caso você as carregue ou importe de um dispositivo, como ao sincronizar uma agenda de contatos.

Se você não usa os Produtos da Meta ou os usa sem uma conta, suas informações ainda podem ser coletadas. Saiba mais sobre como a Meta usa as informações de contato carregadas pelos titulares das contas.

Saiba como carregar e excluir contatos no Facebook e no Messenger ou como conectar a lista de contatos do seu dispositivo no Instagram.

As informações que coletamos ou inferimos sobre você com base na atividade de outras pessoas

Coletamos informações sobre você com base na atividade de outras pessoas. Veja alguns exemplos.

Também inferimos informações sobre você com base na atividade de outras pessoas. Veja dois exemplos:

- Podemos sugerir um amigo pelo recurso Pessoas que você talvez conheça, do Facebook, se os dois aparecem em uma lista de contatos que alguém carregou.
- Levamos em consideração se seus amigos pertencem a um grupo quando sugerimos que você participe dele.³³⁹

Demonstra-se, minimamente, a explicação ao usuário sobre a forma da coleta de dados. Quanto à atualização dos Termos de Serviço, a *Meta* verbaliza a inclusão dos direitos e

³³⁹ FACEBOOK. **Política de Privacidade**. 2022. Disponível em: <https://www.facebook.com/privacy/policy>. Acesso em: 20/07/2022.

obrigações descritos nos nossos Padrões da Comunidade, “como quando podemos desativar ou encerrar contas que violem nossos Termos ou Padrões da Comunidade, os direitos de propriedade intelectual de terceiros ou outras leis”³⁴⁰. Na Seção “4. Disposições adicionais”, no item “2. Suspensão ou encerramento da conta”, dispõe laconicamente o seguinte:

2. Suspensão ou encerramento da conta

Queremos que o Facebook seja um espaço em que as pessoas se sintam bem-vindas e seguras para se expressar e compartilhar seus pensamentos e ideias.

Se nós determinarmos, a nosso critério, que tenha violado clara, séria ou repetidamente nossos Termos ou Políticas, incluindo especialmente nossos Padrões da Comunidade, podemos suspender ou desativar permanentemente seu acesso aos Produtos das Empresas da Meta e podemos desativar ou excluir sua conta permanentemente. Também podemos desativar ou excluir sua conta se você violar recorrentemente os direitos de propriedade intelectual de outra pessoa ou quando formos obrigados por motivos legais.

Podemos desativar ou excluir sua conta se ela não for confirmada após o registro, se sua conta não for utilizada e permanecer inativa por um período prolongado, ou se detectarmos que alguém pode ter usado sua conta sem sua permissão e não conseguirmos confirmar a propriedade da sua conta. Saiba mais sobre como desativamos e excluimos contas.

Se realizarmos essas ações, avisaremos a você e explicaremos suas opções para solicitar uma análise, a menos que isso possa nos expor ou expor outros a responsabilidades legais; prejudicar nossa comunidade de usuários; comprometer ou interferir na integridade ou operação de qualquer um de nossos serviços, fontes ou Produtos; quando formos restritos devido às limitações técnicas; ou quando formos proibidos de fazê-lo por motivos legais.

Saiba mais sobre o que você pode fazer caso sua conta seja desativada e sobre como entrar em contato conosco se acreditar que desativamos sua conta por engano.

Se você excluir ou se nós desativarmos ou excluirmos sua conta, estes Termos serão encerrados como um acordo entre você e nós, mas as seguintes disposições continuarão vigentes: 3 e 4.2 a 4.5.³⁴¹

Ademais, verifica-se que a Política de Privacidade atualizada da *Meta* abrange o *Facebook*, o *Instagram*, o *Messenger* e outros. Contudo, não abrange o *WhatsApp*, o *Workplace*, o *Free Basics*, o *Messenger Kids* ou o uso de dispositivos *Quest*, sem uma conta no *Facebook*, em decorrência de essas redes possuírem suas próprias políticas de privacidade.³⁴²

As alterações de maneira geral, tiveram um propósito de tornar a plataforma mais transparente, trazendo uma linguagem mais prática e clara, proporcionando uma facilitação no entendimento dos diversos tipos e níveis de usuários espalhados pelo mundo. À primeira vista, parece ser um avanço nos Termos de Uso, sobretudo num esforço de aproximação e

³⁴⁰ PROTTI, Michel. O que você precisa saber sobre a atualização da nossa política de privacidade e em nossos termos de serviço. **Meta**. 2022. Disponível em: <https://about.fb.com/br/news/2022/05/o-que-voce-precisa-saber-sobre-a-atualizacao-da-nossa-politica-de-privacidade-e-em-nossos-termos-de-servico/>. Acesso em: 20/06/2022.

³⁴¹ FACEBOOK. **Termos de Serviço**. 2022. Disponível em: https://www.facebook.com/legal/terms/plain_text_terms. Acesso em: 25/06/2022.

³⁴² PROTTI, Michel. O que você precisa saber sobre a atualização da nossa política de privacidade e em nossos termos de serviço. **Meta**. 2022. Disponível em: <https://about.fb.com/br/news/2022/05/o-que-voce-precisa-saber-sobre-a-atualizacao-da-nossa-politica-de-privacidade-e-em-nossos-termos-de-servico/>. Acesso em: 20/06/2022.

transparência com os usuários finais, entretanto, a luta pelo aperfeiçoamento é constante e será mais bem abalizada na prática, a conferir. Colacionamos a figura abaixo, proposta pelo *Facebook* aos usuários:

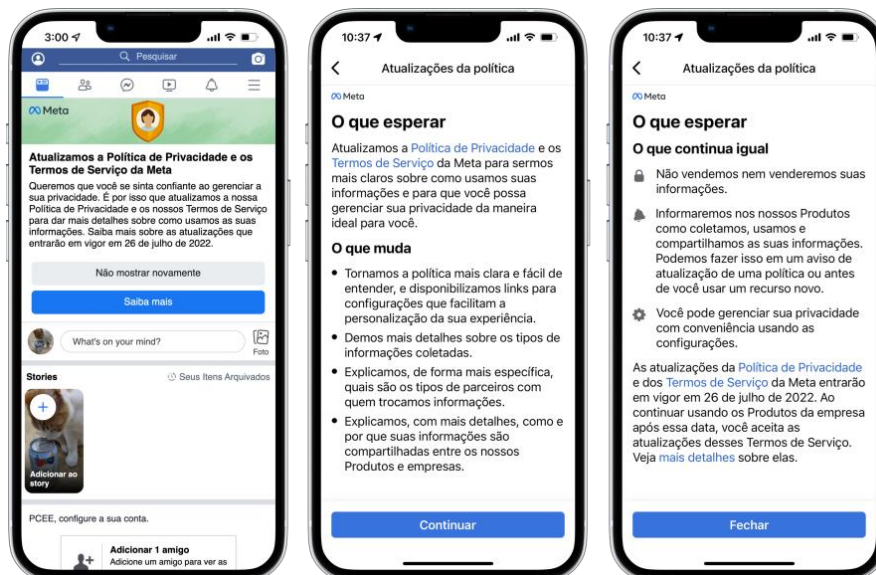


Figura 1: Demonstração da Atualização da Política do *Facebook*

(Fonte: <https://about.fb.com/br/news/2022/05/o-que-voce-precisa-saber-sobre-a-atualizacao-da-nossa-politica-de-privacidade-e-em-nossos-termos-de-servico/>.)

Entretanto, conclui que as alterações, embora demonstrassem um certo avanço, não representam uma solução para ausência de transparência dos processos internos da plataforma, principalmente, em relação a solicitações privadas de remoção de conteúdo.³⁴³ O *Ranking Digital Rights*, publicou um Índice de Responsabilidade Corporativa, contendo 16 empresas de *Internet*, dentre elas, o *Facebook*. Constatou-se, quanto às solicitações privadas de remoção de conteúdo, que os principais intermediários forneceram informações mínimas. Bem como, nenhuma das empresas divulgam informações sobre as ações adotadas para garantir o fiel cumprimento de seus termos de serviço.³⁴⁴

2.4.3. O primeiro relatório sobre direitos humanos da *Meta* e a Declaração para o Futuro da *Internet*: avanços e perspectivas.

³⁴³ LAIDLAW, Emily B. Online Platform Responsibility and Human Rules. In: BELLI, Luca; ZINGALES, Nicolo. **Platform regulations: how platforms are regulated and how they regulate us**. Official outcome of the UN IGF Dynamic Coalition on Platform Responsibility. – Rio de Janeiro: Escola de Direito do Rio de Janeiro da Fundação Getúlio Vargas, 2017. p. 69-70. Disponível em: <https://bibliotecadigital.fgv.br/dspace/handle/10438/19402>. Acesso em: 17/05/2022.

³⁴⁴ RANKING DIGITAL RIGHTS. **2015 Corporate Accountability Index**. 2015. Disponível em: <https://rankingdigitalrights.org/index2015/assets/static/download/RDR-4pager.pdf>. Acesso em: 20/06/2022.

Nesse contexto, de proteção de direitos humanos no ambiente *online*, principalmente a necessidade de sua observância pelos atores privados, numa visão horizontal de aplicação desses direitos³⁴⁵, além da necessidade de novas posturas do Estado, como integrante da governança da *internet*, frente ao desafio de garantidor de uma *Internet* segura, livre, ampla e sem violação dos direitos fundamentais, apresentamos dois esforços recentes, os quais caminham ou demonstram uma tendência em alinhar a atuação do Estado e das empresas privadas de infraestrutura *online* aos ditames dos direitos humanos, tarefa de difícil aplicação, mas urgente e necessária.

O primeiro esforço é a elaboração, em março de 2021, do Primeiro Relatório de Direitos Humanos da *Meta*, empresa que detém várias plataformas, dentre elas, o *Facebook*. O documento é fruto de um compromisso firmado em sua Política de Direitos Humanos,³⁴⁶ comprometendo-se a realizar anualmente um relatório das medidas adotadas sobre os impactos nos direitos humanos, tratando-se do primeiro relatório anual, referente às atividades do período de 1º de janeiro de 2020 a 31 de dezembro de 2021 e sendo baseado no Princípio 15 dos Princípios Orientadores da ONU sobre Empresas e Direitos Humanos, que dispõe:

15. A fim de cumprir sua responsabilidade de respeitar os direitos humanos, as empresas devem ter políticas e processos adequados ao seu tamanho e circunstâncias, incluindo:

- (a) Um compromisso político para cumprir sua responsabilidade de respeitar os direitos humanos;
- (b) Um processo de *due diligence* em direitos humanos para identificar, prevenir, mitigar e prestar contas de como eles abordam seus impactos sobre os direitos humanos;
- (c) Processos para permitir a remediação de quaisquer impactos adversos aos direitos humanos que causem ou para os quais contribuam.³⁴⁷

A partir dessa perspectiva, a *Meta* passou a reconhecer a sua atividade com o potencial de impactar os direitos humanos, passando a afirmar, devido a essa constatação, a adoção de medidas para reparar o seu comportamento diante das premissas fundamentais para o conteúdo constitucional, demonstrando o compromisso à garantia dos direitos humanos.

³⁴⁵ CALLAMARD, Agnès. The Human Rights Obligations of Non-State Actors. In: JØRGENSEN, Rikk Frank [Ed.]. **Human Rights is the Age Platforms**. Cambridge, MA: The MIT Press, 2019. p. 191. Disponível em: <https://direct.mit.edu/books/book/4531/Human-Rights-in-the-Age-of-Platforms>. Acesso em: 29/10/2021.

³⁴⁶ META. **Corporate Human Rights Policy**. Disponível em: <https://about.fb.com/wp-content/uploads/2021/04/Facebooks-Corporate-Human-Rights-Policy.pdf>. Acesso em: 26/06/2022.

³⁴⁷ Tradução livre do original: “15. In order to meet their responsibility to respect human rights, business enterprises should have in place policies and processes appropriate to their size and circumstances, including: (a) A policy commitment to meet their responsibility to respect human rights; (b) A human rights due diligence process to identify, prevent, mitigate and account for how they address their impacts on human rights; (c) Processes to enable the remediation of any adverse human rights impacts they cause or to which they contribute.” (CONSELHO DE DIREITOS HUMANOS DAS NAÇÕES UNIDAS. **Princípios Orientadores sobre Negócios e Direitos Humanos das Nações Unidas**. 2011. Disponível em: https://www.ohchr.org/sites/default/files/documents/publications/guidingprinciplesbusinessshr_en.pdf. Acesso em: 20/01/2022. p. 15-16.)

Assim, explica: “buscamos traduzir a orientação dos direitos humanos em ações significativas, todos os dias.”³⁴⁸

Marechal e Rydzak asseveram esse lançamento como resultado de uma série de críticas feitas pela sociedade civil, a qual apontava falhas da *Meta* em propor respostas às acusações de omissão aos abusos ocorridos no ambiente *online*, cuja repercussão, de forma geral, influenciava no mundo *offline*³⁴⁹. Todavia, mesmo havendo satisfação com o relatório apresentado, esperavam “um maior reconhecimento das políticas e estruturas de incentivo existentes que estão sufocando a capacidade da empresa de responder melhor às questões de direitos humanos.”³⁵⁰

Inobstante, Marechal e Rydzak dialogam sobre a melhor parte do relatório ser a sua existência; esse pensamento surge, pois o “relatório faz referências explícitas às normas e instrumentos internacionais de direitos humanos, é um começo. Antes tarde do que nunca!”³⁵¹ Nesse sentido, verifica-se ponto importante em razão da *Meta* dizer assegurar a promoção à liberdade de expressão à gerência dos riscos relacionados ao equilíbrio da liberdade de expressão e a outros direitos fundamentais nas plataformas de mídia social, bem como:

Durante o período do relatório, atualizamos nossas políticas como parte de nosso esforço contínuo para equilibrar esses direitos em um mundo cada vez mais dinâmico: incluindo, especificamente, fazer referência a princípios de direitos humanos, esclarecer nossas políticas de desinformação de saúde, aprimorar nossa política de bullying e assédio para incluir proteções mais fortes contra assédio baseado em gênero para todos, incluindo figuras públicas, decretando proibições contra certos assédios ou brigadas em massa e expandindo nossas políticas que proíbem ameaças veladas e implícitas, entre outros. Além disso, o Meta fornece aos usuários um caminho para discordar das decisões de política de conteúdo, conforme descrito em mais detalhes neste relatório em *Fornecendo acesso ao remédio*.³⁵²

³⁴⁸ Tradução livre do original: “[...] we seek to translate human rights guidance into meaningful action, every day”. (META. **Meta Human Rights Report: Insights and Actions 2020-2021**. Julho de 2022. p. 4, Disponível em: https://about.fb.com/wp-content/uploads/2022/07/Meta_Human-Rights-Report-July-2022.pdf. Acesso em: 20/07/2022.)

³⁴⁹ MARECHAL, Natalie; RYDZAK, Jan. Meta’s First Human Rights Report: The Good, the Bad, and the Missing. **Ranking Digital Rights**. 2022. Disponível em: <https://rankingdigitalrights.org/2022/07/20/metass-first-human-rights-report-the-good-the-bad-and-the-missing/>. Acesso em: 28/07/2022.

³⁵⁰ MARECHAL, Natalie; RYDZAK, Jan. Meta’s First Human Rights Report: The Good, the Bad, and the Missing. **Ranking Digital Rights**. 2022. Disponível em: <https://rankingdigitalrights.org/2022/07/20/metass-first-human-rights-report-the-good-the-bad-and-the-missing/>. Acesso em: 28/07/2022.

³⁵¹ MARECHAL, Natalie; RYDZAK, Jan. Meta’s First Human Rights Report: The Good, the Bad, and the Missing. **Ranking Digital Rights**. 2022. Disponível em: <https://rankingdigitalrights.org/2022/07/20/metass-first-human-rights-report-the-good-the-bad-and-the-missing/>. Acesso em: 28/07/2022.

³⁵² Tradução livre do original: “During the reporting period, we updated our policies as part of our continuing effort to balance these rights in an increasingly dynamic world: including, specifically referencing human rights principles, clarifying our health misinformation policies, enhancing our bullying and harassment policy to include stronger protections against gender-based harassment for everyone including public figures, enacting prohibitions against certain mass harassment or brigading, and expanding our policies that prohibit veiled and implicit threats, among others. Meta further provides users with a path to disagree with content policy decisions, as described in greater detail in this report under Providing Access to Remedy.” (META. **Meta Human Rights**

Ademais, outro detalhe a ser destacado é o reconhecimento, pela *Meta*, sobre os direitos humanos não se restringirem aos usuários, sendo lógica, ainda assim, sua aplicação aos mesmos; adiante, os impactos causados pelo *Facebook* transpassam os limites da plataforma. Sobre isso, cita-se o documento a respeito dos indivíduos sob a proteção das normas e prerrogativas serem “não apenas usuários de nossas plataformas e serviços, mas muitos outros cujos direitos foram potencialmente impactados pela atividade e conduta online”.³⁵³

Marechal e Rydzak, indicam como ponto positivo a utilização do Programa de Parceiros Confiáveis do *Facebook* – *Trusted Partner* (TP), o qual, segundo o relatório, é um canal de comunicação composto por “mais de 400 organizações não governamentais, agências humanitárias, defensores de direitos humanos e pesquisadores de 113 países ao redor do mundo”³⁵⁴. O objetivo do *software* é auxiliar o *Facebook* a compreender os impactos de suas políticas aos usuários em situação de risco. Contudo, não foram fornecidas informações suficientes sobre o funcionamento do aplicativo, apesar de seu uma prática positiva, peca na transparência para a avaliação realista do programa.³⁵⁵

As críticas aos Relatório, ressalvadas pelo seu recente lançamento, na visão de Marechal e Rydzak, giram entorno das afirmações feitas pela *Meta* as quais, na percepção dos autores, são tendenciosas e não condizem com a realidade. As observações têm base a partir da compreensão do trecho sobre a missão do documento: “[...] dar às pessoas o poder de construir uma comunidade e aproximar o mundo”. De alguma forma, de acordo com *Meta*, essa declaração deve se alinhar “fortemente” com “princípios de direitos humanos”.³⁵⁶

Posteriormente, a *Meta* defende o conceito de “obrigação universal de não discriminação” interligado à sua “visão e estratégia”; Contudo, Marechal e Rydzak asseveram a não admissão sobre a existência do modelo de negócios de publicidade direcionada permitir

Report: Insights and Actions 2020-2021. Julho de 2022. p. 5, Disponível em: https://about.fb.com/wp-content/uploads/2022/07/Meta_Human-Rights-Report-July-2022.pdf. Acesso em: 20/07/2022).

³⁵³ Tradução livre do original: “[...] not only users of our platforms and services, but the many others whose rights were potentially impacted by online activity and conduct.” (META. **Meta Human Rights Report:** Insights and Actions 2020-2021. Julho de 2022. p. 55, Disponível em: https://about.fb.com/wp-content/uploads/2022/07/Meta_Human-Rights-Report-July-2022.pdf. Acesso em: 20/07/2022)

³⁵⁴ Tradução livre do original: “[...] of over 400 key non-governmental, not-for-profit, humanitarian, and international organizations in 113 countries.” (META. **Meta Human Rights Report:** Insights and Actions 2020-2021. Julho de 2022. p. 81, Disponível em: https://about.fb.com/wp-content/uploads/2022/07/Meta_Human-Rights-Report-July-2022.pdf. Acesso em: 20/07/2022)

³⁵⁵ MARECHAL, Natalie; RYDZAK, Jan. Meta’s First Human Rights Report: The Good, the Bad, and the Missing. **Ranking Digital Rights**. 2022. Disponível em: <https://rankingdigitalrights.org/2022/07/20/metass-first-human-rights-report-the-good-the-bad-and-the-missing/>. Acesso em: 28/07/2022.

³⁵⁶ MARECHAL, Natalie; RYDZAK, Jan. Meta’s First Human Rights Report: The Good, the Bad, and the Missing. **Ranking Digital Rights**. 2022. Disponível em: <https://rankingdigitalrights.org/2022/07/20/metass-first-human-rights-report-the-good-the-bad-and-the-missing/>. Acesso em: 28/07/2022.

e automatizar a discriminação, fundamentados em dados demográficos e no comportamento; bem como a sua omissão à fomentação de preconceitos resultantes da desigualdade da *Meta* em alocar recursos para moderação de conteúdo em diferentes idiomas.³⁵⁷

Em suma, a pesquisa apresentou aspectos positivos e negativos, ressaltando-se a positiva iniciativa em fornecer dados; entretanto, é imprescindível a evolução da empresa aos pontos negativos demonstrados acima. Marechal e Rydzak, destacam: (a) Existe pouco tratamento sobre a política de anúncios no relatório: embora a empresa ganhe mais de 98% de seu dinheiro com publicidade, o debate sobre os efeitos do conteúdo e dos sistemas de anúncios da *Meta* quase não aparece em suas “avaliações de impacto nos direitos humanos”. (b) o Relatório indica que a *Meta* criou sistemas classificadores de IA, os quais, segundo eles, permitirão “impor proibições à violação de anúncios e listagens comerciais de certos produtos médicos”. Esta parece ser a única referência a anúncios em todo o relatório. (c) A equipe de políticas de direitos humanos da *Meta*, responsável por este relatório, contava com quatro funcionários, em tempo integral, no final de 2021. Esse grupo de trabalho pareceu pequeno para investigar adequadamente a política de direitos humanos de uma empresa do porte da *Meta*; (d) Não há menção quanto à aplicação desigual das políticas de conteúdo entre regiões, países e idiomas. Consta-se, dessa forma, a menção de uma moderação de conteúdo orientada por IA e uma promessa de “melhorar nossa moderação em todos os idiomas, adicionando mais experiência”, contudo, silencia sobre como isso afeta sua capacidade de moderar efetivamente ou como os direitos humanos são afetados.³⁵⁸

Em síntese, como visto, o referido relatório deixou de tratar de alguns temas, aprofundando melhor, por exemplo, a relação de direitos humanos, ou ainda, uma maior contenção de danos, quando infringidos os limites da liberdade de expressão e outros são abordados insuficientes. Entretanto, representou grande avanço, pois a *Meta* reconheceu a necessidade do comprometimento com uma política de direitos humanos.

É cediço que a grande maioria das empresas de tecnologia não tem a preocupação de elaborar algum relatório como este. Ademais, constatou-se que o relatório é lacônico quanto às políticas de moderação de conteúdo. Com isso, fica lançado a pretensão que a empresa irá

³⁵⁷ MARECHAL, Natalie; RYDZAK, Jan. *Meta's First Human Rights Report: The Good, the Bad, and the Missing*. **Ranking Digital Rights**. 2022. Disponível em: <https://rankingdigitalrights.org/2022/07/20/metas-first-human-rights-report-the-good-the-bad-and-the-missing/>. Acesso em: 28/07/2022.

³⁵⁸ MARECHAL, Natalie; RYDZAK, Jan. *Meta's First Human Rights Report: The Good, the Bad, and the Missing*. **Ranking Digital Rights**. 2022. Disponível em: <https://rankingdigitalrights.org/2022/07/20/metas-first-human-rights-report-the-good-the-bad-and-the-missing/>. Acesso em: 28/07/2022.

aderir aos princípios de direitos humanos. De toda sorte, como afirmam Marechal e Rydzak, o primeiro passo para resolver um problema é admitir que ele existe.³⁵⁹

Mesmo numa vertente diferente, entendemos ser salutar estabelecer a pretensão da *Declaração para o Futuro da Internet*³⁶⁰, atualmente, como mais uma forma de avanço no debate do tema em empresa de mídias sociais e liberdade de expressão, além de representar uma perspectiva de melhora. Tal declaração foi proposta em 28 de abril de 2022, pela União Europeia e Estados Unidos, juntamente com parceiros internacionais, como uma forma de tentar retomar uma postura positiva para a governança da *Internet*.

Mueller lembra ter ocorrido, durante sua formatação da declaração, a alteração da ideia inicial de uma “aliança” para uma “declaração”, com a aliança transparecia uma coalisão entre governos, levando ao entendimento de inclusão de alguns membros em detrimento a outros, os quais poderiam ser excluídos. Com a alteração do termo, o resultado foi uma ampla declaração de princípios, a qual os governos podem simplesmente aderir, adicionando seus nomes, se concordarem. Atualmente a Declaração foi endossada por cerca de 60 países, contudo, ainda existindo países importantes não signatários, como Índia, Chile e Brasil, sendo pauta para o Governo de Joe Biden, num exercício construtivo conseguir convencê-los.³⁶¹

A Declaração possui, como principal objetivo, “recuperar a promessa da *internet*”, além de denunciar as tendências recentes, de governos autoritários, de repressão à liberdade de expressão e negação de direitos humanos e liberdades fundamentais, solicitando aos signatários, além disso, afirmarem o “compromisso de promover e sustentar uma *Internet*”, buscando sempre uma *Internet* “aberta, livre, global, interoperável, confiável e segura”, principalmente para “garantir que a *Internet* reforce os princípios democráticos e os direitos humanos e as liberdades fundamentais”. O acordo contempla oportunidades para pesquisa e comércio colaborativos; estabelece um sistema desenvolvido, governado e implantado de forma inclusiva, para que comunidades não atendidas, ou mal atendidas, especialmente

³⁵⁹ MARECHAL, Natalie; RYDZAK, Jan. *Meta’s First Human Rights Report: The Good, the Bad, and the Missing*. **Ranking Digital Rights**. 2022. Disponível em: <https://rankingdigitalrights.org/2022/07/20/metass-first-human-rights-report-the-good-the-bad-and-the-missing/>. Acesso em: 28/07/2022.

³⁶⁰ **DECLARATION FOR THE FUTURE OF INTERNET**. 2022. Disponível em: https://www.whitehouse.gov/wp-content/uploads/2022/04/Declaration-for-the-Future-for-the-Internet_Launch-Event-Signing-Version_FINAL.pdf. Acesso em: 30/05/2022.

³⁶¹ MUELLER, Milton. *The Declaration for the Future of Internet*. **Blog of the Internet Governance Project (IGP)**, 29/04/2022. Disponível em: <https://www.internetgovernance.org/2022/04/29/the-declaration-for-the-future-of-the-internet/>. Acesso em 30/05/2022.

aquelas que estão *online* pela primeira vez, possam navegar com segurança, privacidade e proteções de dados pessoais em vigor; e é regido por processos multissetoriais.³⁶²

Daniele Duca, pontua que essa visão da Declaração, no entanto, parece não ser tão diferente daquela já estabelecida com o Mercado Único Digital e observando os pilares da Declaração, percebe-se o crescimento mais lento do que o esperado,³⁶³ sendo os compromissos sugeridos inseridos nas seguintes áreas: Direitos humanos e liberdades fundamentais; Compatibilidade e governança globais; Acesso inclusivo e acessível; Confie no ecossistema digital; e Governança multissetorial.

Percebe-se, a iniciativa limitada do documento, porém, tal acordo útil como reação apoiada pelo Estado contra uma escalada autoritarista, tanto na repressão de liberdade, como na economia política digital.

Portanto, nesse quadro preocupante, reside a questão primordial sobre o alcance dos remédios constitucionais, limitantes das ameaças ao princípio do Estado de Direito, na sociedade da informação, sendo crucial limitar o exercício de poderes públicos e privados autônomos, os quais prejudicam a proteção dos direitos e liberdades fundamentais. Oreste Pollicino aponta como uma das saídas possíveis para a problemática, a aplicação horizontal dos direitos fundamentais em relação a particulares.³⁶⁴

2.5. A eficácia horizontal dos direitos fundamentais na proteção da liberdade de expressão nas plataformas de mídia social.

O debate sobre a eficácia dos direitos fundamentais no âmbito das relações entre particulares, teve seu ápice no direito alemão, entre as décadas de 1950 e 1960. Nos Estados Unidos, ficou conhecido como *State Action*, em que a aplicação dos direitos fundamentais às relações privadas passou a ser discutida. Mendes e Fernandes, argumentam que após esse debate, o reconhecimento da dimensão horizontal dos direitos fundamentais pela via da

³⁶² MUELLER, Milton. The Declaration for the Future of Internet. **Blog of the Internet Governance Project (IGP)**, 29/04/2022. Disponível em: <https://www.internetgovernance.org/2022/04/29/the-declaration-for-the-future-of-the-internet/>. Acesso em 30/05/2022.

³⁶³ Tradução livre do original: “The vision, though, appears not to be that different from the one that was already established with the Digital Single Market. Going over the pillars of the Declaration, it is clear that the growth was slower than expected”. (DUCA, Danielle. **Internet in the European Union: past, present and future of digitalization**. 2021. 118 f. Dissertação (Mestrado em Estudos Europeus e Globais). Università Degli Studi di Padova. Pádua, 2021. p. 90. Disponível em: https://thesis.unipd.it/bitstream/20.500.12608/32989/1/Daniele_Duca.pdf. Acesso em 30/06/2022).

³⁶⁴ POLLICINO, Oreste. **Digital Private Powers Exercising Public Functions: The Constitutional Paradox in the Digital Age and its Possible Solutions**. 2021. Disponível em: https://echr.coe.int/Documents/Intervention_20210415_Pollicino_Rule_of_Law_ENG.pdf. Acesso em: 27/10/2021.

jurisdição constitucional passou a tomar a forma da aferição do papel do legislador diante da perspectiva de dever de proteção. Desta forma, Cortes Constitucionais passaram a avaliar o cumprimento, por parte do legislador, não apenas uma proibição do excesso, mas também uma proibição de omissão, na adequação dos direitos fundamentais na dimensão privada.³⁶⁵

Como demonstrado, as plataformas de mídia social assumem posturas quase governamental ao estipularem normas, ditando o ritmo das discussões, determinado – muitas vezes – o tema, como e quando ocorrerá os debates, não recaindo tais escolhas somente sobre posicionamentos considerados ilegais, mas também, sobre as manifestações legais que sejam inconvenientes ao modelo de negócio, despertando, notadamente, vários questionamentos, dentre os quais, o que discutimos neste último item do capítulo, o qual, de forma sintética, aborda dentro da teoria dos direitos fundamentais, a parte da eficácia nas relações entre particulares, com ênfase na relação entre plataformas *online* e usuários.³⁶⁶

Atualmente, a doutrina majoritária, quanto a possibilidade de aplicação de direitos fundamentais nas relações entre particulares, é dizer, que os direitos fundamentais exerçam sua eficácia vinculante também na esfera jurídico-privada³⁶⁷ e, nesse sentido, fazem parte Sarlet³⁶⁸, Canotilho³⁶⁹, Sarmiento³⁷⁰, Branco³⁷¹, dentre outros. Como aduz Pollicino, comentando a suposição de Robert Alexy, de que a questão do efeito horizontal dos direitos fundamentais não pode ser desvinculada em termos teóricos da questão mais geral do efeito

³⁶⁵ MENDES, Gilmar Ferreira; FERNANDES, Victor Oliveira. Eficácia dos direitos fundamentais nas relações privadas da internet: o dilema da moderação de conteúdo em redes sociais na perspectiva comparada Brasil-Alemanha. **Revista de Direito Civil Contemporâneo**. v. 31. ano 9. p. 33-68. São Paulo: Ed. RT, abr./jun. 2022. p. 2. Disponível em: <http://revistadostribunais.com.br/maf/app/document?stid=st-rql&marg=DTR-2022-9857>. Acesso em: 05/08/2022.

³⁶⁶ Nesse sentido, Sarlet e Hartmann afirmam que “é possível perceber o peso da atuação na internet de entes privados sobre o exercício de direitos fundamentais como a liberdade de expressão, o acesso à informação, a privacidade, a honra e a propriedade intelectual. Plataformas privadas atuam nesse cenário como entes estatais que definem e aplicam unilateralmente as regras sobre o que pode ou não ser dito, efetivamente influenciando os limites da liberdade de expressão em diferentes países muito mais do que a legislação ou o Judiciário local. Adicione-se a isso o fato de que essa governança é feita com pouca ou nenhuma transparência.” (SARLET, Ingo Wolfgang; HARMAN, Ivar Alberto Martins. *Direitos Fundamentais e Direito Privado: a Proteção da Liberdade de Expressão nas Mídias Sociais*. **Revista Direito Público**. Porto Alegre, v. 16, n. 90, p. 85-108, 2019. p. 97. Disponível em: https://repositorio.pucrs.br/dspace/bitstream/10923/18863/2/Direitos_Fundamentais_e_Direito_Privado_a_Protecao_da_Liberdade_de_Expressao_nas_Mdias_Sociais.pdf. Acesso em: 24/10/2021).

³⁶⁷ SARLET, Ingo Wolfgang. **A eficácia dos direitos fundamentais: uma teoria geral dos direitos fundamentais na perspectiva constitucional**. 10 ed. 2. Tir. Porto Alegre: Livraria do Advogado Editora, 2010. p. 374.

³⁶⁸ SARLET, Ingo Wolfgang. **A eficácia dos direitos fundamentais: uma teoria geral dos direitos fundamentais na perspectiva constitucional**. 10 ed. 2. Tir. Porto Alegre: Livraria do Advogado Editora, 2010. p. 375-383.

³⁶⁹ CANOTILLO, J. J. GOMES. **Direito constitucional e teoria da constituição**. 7ª ed., 21 reimp. Coimbra: Almedina, 2019. p. 448.

³⁷⁰ SARMENTO, Daniel. **Direitos Fundamentais e Relações Privadas**. Rio de Janeiro: Lumen Juris. 2004. p.

³⁷¹ BRANCO, Paulo Gustavo Gonet. Teoria Geral dos Direitos Fundamentais. In: MENDES, Gilmar Ferreira; BRANCO, Paulo Gustavo Gonet. **Curso de direito constitucional**. 15. ed., São Paulo: Saraiva, 2020. p. 177-183.

direto dos mesmos direitos. Pollicino argumenta, ainda, que uma vez reconhecido o efeito direto de um direito fundamental, este passa a ter uma dupla dimensão, sendo a primeira vertical, no sentido da relação clássica 'autoridade pública vs liberdade individual', e a segunda dimensão horizontal, concentrada na relação entre privados, bem como na relação menos clássica, entre novos poderes privados e indivíduos/usuários.³⁷²

Hoffmann-Riem, acrescenta, nessa linha de raciocínio, o fato de, para este efeito horizontal dos direitos fundamentais, os seguintes fatores desempenharem papel decisivo: “o desequilíbrio entre as partes contrárias, o significado social de certos serviços ou o poder social de um dos lados.”³⁷³

Nessa relação, entre as plataformas *online* e os usuários, nos apoiamos na concepção – clássica – defendida por Sarlet: “é preciso reconhecer uma eficácia direta *prima facie* dos direitos fundamentais, especialmente da liberdade de expressão, nas relações entre plataformas digitais privadas e seus usuários.”³⁷⁴ Contudo, nesse ponto, reside uma certa resistência da doutrina³⁷⁵ – minoritária – em adotar a eficácia direta ou eficácia imediata.

A Constituição Federal de 1998 dispõe expressamente no Art. 5º, § 1º, que as normas definidoras dos direitos e garantias fundamentais possuem aplicação imediata, o que, ao nosso ver – e da doutrina majoritária³⁷⁶ – demonstra ser uma decisão patente em reconhecer a eficácia direta das normas de direitos fundamentais, ou seja, “todos os órgãos estatais estão obrigados a assegurar a maior efetividade e proteção possível aos direitos fundamentais”.³⁷⁷ Acrescente-se a definição de Alexy sobre eficácia direta ou imediata perante terceiros, a qual consiste em que “[...] por razões ligadas aos direitos fundamentais, há determinados direitos e não-direitos, liberdades e não-liberdades e competências e não-competências na relação

³⁷² POLLICINO, Oreste. **Digital Private Powers Exercising Public Functions: The Constitutional Paradox in the Digital Age and its Possible Solutions.** 2021. Disponível em: https://echr.coe.int/Documents/Intervention_20210415_Pollicino_Rule_of_Law_ENG.pdf. Acesso em: 27/10/2021.

³⁷³ HOFFMANN-RIEM, Wolfgang. **Teoria Geral do Direito Digital: transformação digital: desafios para o direito.** 2. ed. Rio de Janeiro: Forense, 2022. p. 49.

³⁷⁴ SARLET, Ingo Wolfgang; HARMAN, Ivar Alberto Martins. Direitos Fundamentais e Direito Privado: a Proteção da Liberdade de Expressão nas Mídias Sociais. **Revista Direito Público.** Porto Alegre, v. 16, n. 90, p. 85-108, 2019. p. 98. Disponível em: https://repositorio.pucrs.br/dspace/bitstream/10923/18863/2/Direitos_Fundamentais_e_Direito_Privado_a_Protecao_da_Liberdade_de_Expressao_nas_Mdias_Sociais.pdf. Acesso em: 24/10/2021.

³⁷⁵ Por todos, conferir: DUQUE, Marcelo Schenk. **Direito privado e constituição: drittwirkung** dos direitos fundamentais. São Paulo: Revista dos Tribunais, 2013.

³⁷⁶ Constata-se, também, maioria jurisprudencial, o Supremo Tribunal Federal ao julgar Recurso Extraordinário n. 201818/RJ, da relatoria para o acórdão o Ministro Gilmar Mendes, adotou a tese de uma eficácia direta dos direitos fundamentais nas relações privadas. O caso em julgamento, versava sobre a aplicação da ampla defesa e do contraditório, no afastamento de um sócio de uma sociedade civil, (entidade privada).

³⁷⁷ SARLET, Ingo Wolfgang. **A eficácia dos direitos fundamentais: uma teoria geral dos direitos fundamentais na perspectiva constitucional.** 10 ed. 2. Tir. Porto Alegre: Livraria do Advogado Editora, 2010. p. 270.

cidadão/cidadão, os quais não existiriam sem essas razões.”³⁷⁸ Define-se, portanto, como os efeitos diretos decorrem tanto da teoria dos efeitos indiretos, quanto da teoria dos efeitos mediatos pelo Estado.

Daniel Sarmento, ao analisar a prevalência da eficácia indireta ou direta, defende a aplicação da eficácia direta ao caso concreto. Com tal critério, tornou-se útil no caso da moderação da liberdade de expressão, realizada por plataformas de infraestrutura privada *online*, ante à flagrante *desigualdade material*,³⁷⁹ pois, conforme já explicitado percebe-se a diferença de poder entre a empresa privada de mídia social – como o *Facebook* – e o usuário final³⁸⁰. E resume:

Um dos papéis essenciais dos direitos fundamentais é a proteção da pessoa humana contra o poder. Por isso, como já destacado antes, uma das razões para a extensão destes direitos às relações entre particulares é exatamente a constatação empírica de que, na sociedade contemporânea, existem inúmeros outros polos de poder além do Estado, que podem oprimir o indivíduo.³⁸¹

Nessa esteira, Sarlet pontua que decorre do dever de atribuir máxima efetividade e eficácia às normas de direitos fundamentais, o dever de ser entendido como um mandado de otimização, embora a eficácia e a efetividade dos direitos fundamentais, de maneira geral – e não somente no âmbito das relações entre atores privados – não está disposta a uma lógica do conhecido tipo “tudo ou nada”.³⁸² Propõe, ainda, esse mandado de otimização para enfrentar a “vinculação dos atores da *internet* aos direitos fundamentais”.³⁸³

Nessa concepção da *internet*, conforme apontado por Lessing, foi construída por engenheiros e acadêmicos, contrários ao controle, seguindo o objetivo do exército em criar

³⁷⁸ ALEXY, Robert. **Teoria dos Direitos Fundamentais**. Tradução Virgílio Afonso da Silva. 2 ed. São Paulo: Malheiros Editores, 2011. p. 539.

³⁷⁹ Nesse ponto, é salutar a transcrição do pensamento de Sarmento: “A *desigualdade material* justifica a ampliação da proteção dos direitos fundamentais na esfera privada, porque se parte da premissa de que a assimetria do poder prejudica o exercício da autonomia privada das partes mais débeis. O hipossuficiente, no mais das vezes vai acabar curvando-se diante do arbítrio do mais poderoso, ainda que, do ponto de vista puramente formal, seu comportamento possa parecer decorrente do exercício da sua autonomia privada.” (SARMENTO, Daniel. **Direitos fundamentais e relações privadas**. 2. ed. Rio de Janeiro: Lumen Juris, 2010. p. 262)

³⁸⁰ SARMENTO, Daniel. **Direitos fundamentais e relações privadas**. 2. ed. Rio de Janeiro: Lumen Juris, 2010. p. 264.

³⁸¹ SARMENTO, Daniel. **Direitos fundamentais e relações privadas**. 2. ed. Rio de Janeiro: Lumen Juris, 2010. p. 264.

³⁸² SARLET, Ingo Wolfgang. Neoconstitucionalismo e influência dos direitos fundamentais no direito privado: algumas notas sobre a evolução brasileira. **civilistica.com**, v. 1, n. 1, p. 1-30, 2012. p. 20. Disponível em: <https://civilistica.emnuvens.com.br/redc/article/view/29/17>. Acesso em: 15/06/2022.

³⁸³ SARLET, Ingo Wolfgang; HARMMAN, Ivar Alberto Martins. Direitos Fundamentais e Direito Privado: a Proteção da Liberdade de Expressão nas Mídias Sociais. **Revista Direito Público**. Porto Alegre, v. 16, n. 90, p. 85-108, 2019. p. 93. Disponível em: https://repositorio.pucrs.br/dspace/bitstream/10923/18863/2/Direitos_Fundamentais_e_Direito_Privado_a_Protecao_da_Liberdade_de_Expressao_nas_Mdias_Sociais.pdf. Acesso em: 24/10/2021.

uma rede descentralizada.³⁸⁴ O código foi estruturado para obter – nos primeiros anos de popularidade do *word wide web* – um ambiente leve e democrático. Todavia, com a ascensão econômica e comercial de grandes empresas de mídia social, o código passa a ser alterado para permitir certos privilégios de determinados entes, com mais recursos financeiros.

Como meio de equacionar, Ingo Sarlet e Ivar Hartmman propõem um novo pensamento, sobre qual tipo de eficácia seria mais adequado para ser aplicado nas relações entre particulares nesse contexto. E anotam, que o melhor caminho no momento seria a aplicação simultânea da eficácia indireta como regra e da eficácia direta em casos específicos, numa convivência dialógica.³⁸⁵ Isto é, aplicar a teoria dos deveres de proteção de Claus Wilhelm-Canaris,³⁸⁶ designando ao legislador a atribuição de produzir normas específicas para garantir a proteção necessária dos direitos fundamentais dos autorreguladores – empresas de mídias sociais –, e dos sujeitos dos efeitos dessa autorregulação – usuários (finais) –, devendo ser respeitado os limites da proibição de proteção excessiva e proibição de proteção insuficiente, através de deveres estatais de proteção por meio de normas procedimentais e organizatórias.³⁸⁷

Nessa visão dialógica, por vezes, será preciso aplicar a eficácia direta quando atores privados – plataforma de redes sociais – assumem posição dissonante em sua relação com os usuários, gerando com isso abusos que violem os direitos fundamentais como a liberdade de expressão, além de outros. Essa aplicação será ainda mais necessária, quando envolver violações da dignidade da pessoa humana. Sarlet e Hartmman, advertem que essa aplicação

³⁸⁴ Conferir: LESSIG, Lawrence. **Code**. Version 2.0. New York: Basic Books, 2006. Disponível em: https://cyber.harvard.edu/ptc2010/sites/ptc2010/images/Lessig_Code_Excerpts.pdf. Acesso em: 04/12/2021.

³⁸⁵ SARLET, Ingo Wolfgang; HARMAN, Ivar Alberto Martins. Direitos Fundamentais e Direito Privado: a Proteção da Liberdade de Expressão nas Mídias Sociais. **Revista Direito Público**. Porto Alegre, v. 16, n. 90, p. 85-108, 2019. p. 93. Disponível em: https://repositorio.pucrs.br/dspace/bitstream/10923/18863/2/Direitos_Fundamentais_e_Direito_Privado_a_Protecao_da_Liberdade_de_Expressao_nas_Mdias_Sociais.pdf. Acesso em: 24/10/2021.

³⁸⁶ “As proibições de discriminação fundadas na dignidade humana e na sua proteção vinculam diretamente (no âmbito de uma eficácia em relação a terceiros ou eficácia privada) os atores privados. No mais, é reservada ao legislador ‘ordinário’ a possibilidade de estabelecer proibições de discriminação e formular os respectivos suportes fáticos e efeitos jurídicos.” (CANARIS, Claus-Wilhelm. Considerações a respeito da posição de proibições de discriminação no sistema do direito privado. **Revista Brasileira De Direitos Fundamentais & Justiça**, v. 7, n. 22, p. 15–20, 2013. p. 16-17. Disponível em: <https://doi.org/10.30899/dfj.v7i22.279>. Acesso em: 10/12/2021).

³⁸⁷ SARLET, Ingo Wolfgang; HARMAN, Ivar Alberto Martins. Direitos Fundamentais e Direito Privado: a Proteção da Liberdade de Expressão nas Mídias Sociais. **Revista Direito Público**. Porto Alegre, v. 16, n. 90, p. 85-108, 2019. p. 103. Disponível em: https://repositorio.pucrs.br/dspace/bitstream/10923/18863/2/Direitos_Fundamentais_e_Direito_Privado_a_Protecao_da_Liberdade_de_Expressao_nas_Mdias_Sociais.pdf. Acesso em: 24/10/2021.

diferenciada da eficácia entre os diversos atores privados, não significa que não possa haver assimetria, desde que não reste ameaçado os direitos fundamentais.³⁸⁸

Segundo a ideia dos autores, tal postura não indica a necessidade de uma revisão judicial a cada ação nas redes sociais, defendendo essa ação de moderação deve continuar a cargo os atores privados, contudo, enveredando pelo caminho autorregulação regulada – não desenvolveremos esse tema, por exceder os limites deste estudo.³⁸⁹

Oreste Pollicino, sugere um avanço, para além do debate sobre os efeitos horizontais e verticais dos direitos fundamentais no ambiente digital, propondo uma solução alternativa que denomina *habeas corpus* digital de direitos substantivos e processuais. Assim, explica:

Enquanto os direitos substantivos dizem respeito à condição dos indivíduos como sujeitos de um tipo de poder soberano que não é mais exclusivamente exercido pelas autoridades públicas, os direitos processuais decorrem da expectativa de que os indivíduos tenham que reivindicar e fazer valer seus direitos perante outros órgãos que não os órgãos jurisdicionais tradicionais, que empregam métodos diferentes da discricionariedade judicial, como o devido processo tecnológico e horizontal.

Por fim, seguimos as lições de Belli, Francisco e Zingales, para fincar com esperança, expectativas “na autoconsciência das plataformas e na aceitação de sua própria responsabilidade de respeitar os direitos humanos”,³⁹⁰ em harmonia com o reconhecimento das corporações de defesa dos valores dos Direitos Humanos, a fim de “as plataformas não se limitam a conceber as soluções mais eficientes em termos de custos para os conflitos entre os usuários, mas se esforçam para garantir a proteção efetiva dos direitos fundamentais.”³⁹¹ Esperamos a permanência de os Estados “vigilantes para que isso não continue sendo uma

³⁸⁸ SARLET, Ingo Wolfgang; HARMAN, Ivar Alberto Martins. Direitos Fundamentais e Direito Privado: a Proteção da Liberdade de Expressão nas Mídias Sociais. **Revista Direito Público**. Porto Alegre, v. 16, n. 90, p. 85-108, 2019. p. 103. Disponível em: https://repositorio.pucrs.br/dspace/bitstream/10923/18863/2/Direitos_Fundamentais_e_Direito_Privado_a_Protecao_da_Liberdade_de_Expressao_nas_Mdias_Sociais.pdf. Acesso em: 24/10/2021.

³⁸⁹ SARLET, Ingo Wolfgang; HARMAN, Ivar Alberto Martins. Direitos Fundamentais e Direito Privado: a Proteção da Liberdade de Expressão nas Mídias Sociais. **Revista Direito Público**. Porto Alegre, v. 16, n. 90, p. 85-108, 2019. p. 103. Disponível em: https://repositorio.pucrs.br/dspace/bitstream/10923/18863/2/Direitos_Fundamentais_e_Direito_Privado_a_Protecao_da_Liberdade_de_Expressao_nas_Mdias_Sociais.pdf. Acesso em: 24/10/2021.

³⁹⁰ Tradução livre do original: “[...] on platforms’ self-awareness and acceptance of their own responsibility to respect human Rights [...]”. (BELLI, Luca; FRANCISCO, Pedro Augusto; ZINGALES, Nicolo. Law of the Land or Law of the Platform? Beware of the Privatisation of Regulation and Police. p. 41-64. In. BELLI, Luca; ZINGALES, Nicolo. **Platform regulations: how platforms are regulated and how they regulate us**. Official outcome of the UN IGF Dynamic Coalition on Platform Responsibility. Rio de Janeiro: Escola de Direito do Rio de Janeiro da Fundação Getúlio Vargas, 2017.p. 61)

³⁹¹ Tradução livre do original: “[...] platforms do not merely devise the most cost-efficient solutions to conflicts between users, but rather strive to ensure effective protection of fundamental Rights [...]”. (BELLI, Luca; FRANCISCO, Pedro Augusto; ZINGALES, Nicolo. Law of the Land or Law of the Platform? Beware of the Privatisation of Regulation and Police. p. 41-64. In. BELLI, Luca; ZINGALES, Nicolo. **Platform regulations: how platforms are regulated and how they regulate us**. Official outcome of the UN IGF Dynamic Coalition on Platform Responsibility. Rio de Janeiro: Escola de Direito do Rio de Janeiro da Fundação Getúlio Vargas, 2017.p. 61)

aspiração de princípio, garantindo que seja concretizado por meio de políticas de plataformas e ToS - Termos de Uso.”³⁹²

³⁹² Tradução livre do original: “[...] vigilant that this does not remain an aspiration of principle, ensuring that it be given concrete effect through platforms’ policies and ToS.” (BELLI, Luca; FRANCISCO, Pedro Augusto; ZINGALES, Nicolo. Law of the Land or Law of the Platform? Beware of the Privatisation of Regulation and Police. p. 41-64. In. BELLI, Luca; ZINGALES, Nicolo. **Platform regulations**: how platforms are regulated and how they regulate us. Official outcome of the UN IGF Dynamic Coalition on Platform Responsibility. Rio de Janeiro: Escola de Direito do Rio de Janeiro da Fundação Getúlio Vargas, 2017.p. 61)

CAPÍTULO III – MODERAÇÃO DE CONTEÚDO: PARA ALÉM DA REMOÇÃO

Em 6 de janeiro de 2021, uma multidão atacou e invadiu o Capitólio dos Estados Unidos, causando cinco mortes e muitos feridos em um ato de nítida ameaça a ordem constitucional.

O então presidente dos Estados Unidos Donald Trump, realizou duas postagens no *Instagram* e no *Facebook*, uma durante e outra imediatamente após o ataque. Uma das postagens às 16h21 – horário padrão do leste, EUA – em vídeo acusando a eleição presidencial de 2020 de ser “roubada” e “fraudulenta”³⁹³ e a outra postagem era uma declaração às 18h07 – horário padrão do leste, EUA – afirmando: “Estas são as coisas e eventos que acontecem quando uma vitória eleitoral sagrada é tão sem cerimônia e cruelmente despojada de grandes patriotas que foram mal e injustamente tratados por tanto tempo.”³⁹⁴

O *Facebook* – ainda não tinha sido alterado para *Meta* – removeu a primeira postagem às 17h41, horário padrão do leste dos EUA e, a segunda postagem, às 18h15, horário padrão do leste dos EUA – bloqueou as contas de Trump, proibindo-o de realizar novas postagens por período inicial de 24 horas, estendendo o bloqueio indefinidamente, por, pelo menos, duas semanas até a transição do mandato ocorrer, com a posse do presidente eleito Joe Biden. O *Facebook* encaminhou sua decisão de suspender, indefinidamente, as contas de Trump ao seu Conselho de Supervisão, também para apreciação sobre eventuais recomendações do Conselho sobre a suspensão das contas de líderes políticos.³⁹⁵

Registre-se, ainda, em 6 de março de 2020, o *Facebook* anunciou a remoção dos anúncios de máscaras médicas em suas plataformas para impedir que as pessoas tentassem explorar a emergência de saúde pública do COVID-19 para obter ganhos comerciais.³⁹⁶ Cerca de um mês depois, o *New York Times* noticiou que essa proibição do *Facebook*, estava gerando impactos, impedindo voluntários de confeccionarem máscaras costuradas à mão para

³⁹³ UNITED STATES. **Facebook Oversight Board**. Case Decision 2021-001-FB-FBR, OVERSIGHT BD. 9–10. May 5, 2021. Disponível em: <https://www.oversightboard.com/sr/decision/2021/001/pdf-english>. Acesso em: 01/07/2022.

³⁹⁴ Tradução livre do original: “*These are the things and events that happen when a sacred landslide election victory is so unceremoniously & viciously stripped away from great patriots who have been badly & unfairly treated for so long. Go home with love & in peace. Remember this day forever!*” (UNITED STATES. **Facebook Oversight Board**. Case Decision 2021-001-FB-FBR, OVERSIGHT BD. 9–10. May 5, 2021. Disponível em: <https://www.oversightboard.com/sr/decision/2021/001/pdf-english>. Acesso em: 01/07/2022).

³⁹⁵ UNITED STATES. **Facebook Oversight Board**. Case Decision 2021-001-FB-FBR, OVERSIGHT BD. 9–10. May 5, 2021. Disponível em: <https://www.oversightboard.com/sr/decision/2021/001/pdf-english>. Acesso em: 01/07/2022.

³⁹⁶ ROSEN, Guy. An Update on Our Work to Keep People Informed and Limit Misinformation About COVID-19. *Meta*. 2020. Disponível em: <https://about.fb.com/news/2020/04/covid-19-misinfo-update/>. Acesso em: 01/07/2022.

profissionais de saúde, pois os sistemas automatizados de moderação de conteúdo do *Facebook* estavam bloqueando os anúncios de máscara.³⁹⁷ Por outro lado, surgiu a notícia que o *Facebook* estava lucrando com fraudes veiculando anúncios de máscara não capturados por esses mesmos sistemas.³⁹⁸

Posteriormente, em 10 de junho de 2020, o *Facebook* constatou a evolução em relação ao uso de máscaras, desde o início da pandemia, segundo orientações das autoridades de saúde. A proibição, contudo, seria reduzida para permitir a promoção de máscaras não médicas. Entretanto, essa alteração do *Facebook*, como recorda Douek³⁹⁹, somente ocorreu muito tempo depois de os especialistas começarem a recomendar máscaras, e, apenas pouco depois de a OMS mudar suas orientações.⁴⁰⁰ Acrescente-se, ainda, outro evento, ocorrido em fevereiro de 2021, em que o *Facebook* proibiu militares de Mianmar usarem suas plataformas do *Facebook* e *Instagram* como efeito imediato, após o golpe aplicado no país sudeste asiático e tomaram o poder.⁴⁰¹

No Brasil, em 10 de agosto de 2022, o *YouTube* removeu de sua plataforma uma *live* do evento realizado no dia 18 de julho de 2022, no qual o presidente Jair Bolsonaro (PL) se reuniu com embaixadores no Palácio do Planalto. Durante o evento, o chefe do executivo fez ataques às urnas eletrônicas, bem como colocou em dúvida o processo eleitoral brasileiro sem apresentar provas; além disso, fez críticas ao Supremo Tribunal Federal e ao Tribunal

³⁹⁷ ISAAC, Mike. Facebook Hampers Do-It-Yourself Mask Efforts, **New York Times**. 2020. Disponível em: <https://www.nytimes.com/2020/04/05/technology/coronavirus-facebook-masks.html>. Acesso em: 01/07/2022.

³⁹⁸ SILVERMAN, Craig. Facebook Banned Mask Ads. They're Still Running. **BuzzFeed News** (May 13, 2020), Disponível em: <https://www.buzzfeednews.com/article/craigsilverman/facebook-mask-ads-ban-zestads-coronavirus>. Acesso em: 01/07/2022.

³⁹⁹ DOUEK, Evelyn. Governing online speech: from “posts-as-trumps” to proportionality and probability. **Columbia Law Review**. v. 121, n. 3, p. 759-837, 2021. p. 761-762 Disponível em: https://www.columbialawreview.org/wp-content/uploads/2021/04/Douek-Governing_Online_Speech-from_Posts_As-Trumps_To_Proportionality_And_Probability.pdf. Acesso em: 03/11/2021.

⁴⁰⁰. Oreste Pollicino, acrescenta ainda, o episódio que o Facebook proibiu editores e usuários australianos de compartilhar ou visualizar conteúdo de notícias australianas e internacionais. Levando o Governo Australiano a negociar com os editores de notícias, pagar por notícias, compartilhar dados e receitas de publicidade. Além disso, o pedido do governo da Índia ao Twitter para bloquear mais de mil contas de apoio aos protestos de agricultores levou as mídias sociais a cumprir essa ordem pública. (POLLICINO, Oreste. **Digital Private Powers Exercising Public Functions: The Constitutional Paradox in the Digital Age and its Possible Solutions**. 2021. Disponível em: https://echr.coe.int/Documents/Intervention_20210415_Pollicino_Rule_of_Law_ENG.pdf. Acesso em: 27/10/2021). Consultar a respeito: FARAH, André. **Liberdade de expressão e remoção de conteúdo da internet**. 2ª ed. Rio de Janeiro: Lumen Juris, 2021. p. 121 e seguintes.

⁴⁰¹ POTKIN, Fanny. Facebook bans Myanmar military from its platforms with immediate effect. **Reuters**. 2021. Disponível em: <https://www.reuters.com/article/us-myanmar-politics-facebook-idUSKBN2AP0BK>. Acesso em: 01/07/2022. Para melhor compreensão de como o Facebook, nos anos de 2018-2020, assumiu o papel de árbitro nos conflitos internos de Mianmar, decidindo quais informações eram permitidas e não permitidas, conferir: TØNNESSON, Stein; OO. Min Zaw; AUNG, Ne Lynn. Pretending to be States: The Use of Facebook by Armed Groups in Myanmar. **Journal of Contemporary Asia**. v. 52, n. 2, p. 200-225, 2022. Disponível em: <https://www.tandfonline.com/doi/pdf/10.1080/00472336.2021.1905865>. Acesso em: 01/07/2022.

Superior Eleitoral. O *YouTube* informou que a remoção ocorreu devido à atualização nas diretrizes de combate à desinformação sobre processo eleitoral no *YouTube*.⁴⁰²

Esses eventos sequenciais possuem vários ângulos passíveis de análise; em primeiro momento, a decisão da remoção não está no rol de atividades abarcadas pela liberdade econômica e, seguindo o modelo de negócio das plataformas, como demonstrado no capítulo anterior, essas empresas de mídia social – dentre elas o *Facebook* – consolidaram seus papéis de “guardiões da informação globalmente e dos propósitos comerciais da plataforma”,⁴⁰³ exercendo, desta forma, uma governança privada do discurso *online*, a qual também pode ser denominada como ‘regulação’. Conforme sugere Andrews Leighton, “eu uso o termo ‘regulação’ em vez do termo ‘governança’ popularizado na era das redes,”⁴⁰⁴ justificando essa opção devido “não apenas de seu uso comum no discurso político, onde há uma crença de que estas são coisas até certo ponto diferentes, mas também de um desejo de restabelecer o papel do Estado na promoção do poder compensatório para alcançar valor público”.⁴⁰⁵

Neste capítulo, iniciado pelo caso “*Trump Ban*” – como denominado por Chinmayi Arun⁴⁰⁶, o qual, traduzido para o português, de forma aproximada, seria: *A proibição do Trump* –, serve para lançar à luz o complexo sistema de moderação de conteúdo – especificamente do *Facebook* – e o desempenho de seu Conselho de Supervisão – *Facebook Oversight Board* (FOB). Esse mecanismo de controle estipula a remoção de conteúdo, uma das formas de governança⁴⁰⁷ privada do debate *online*. Sobre esse prisma, esse capítulo objetiva apresentar remédios⁴⁰⁸ alternativos quanto à remoção de conteúdo, que transpasse a

⁴⁰² NIKLAS, Jan. YouTube remove livre de Bolsonaro com embaixadores. **Yahoo.com**. 2022. Disponível em: <https://esportes.yahoo.com/youtube-remove-live-bolsonaro-com-210850760.html>. Acesso em 11/08/2022.

⁴⁰³ POLLICINO, Oreste. **Digital Private Powers Exercising Public Functions: The Constitutional Paradox in the Digital Age and its Possible Solutions**. 2021. Disponível em: https://echr.coe.int/Documents/Intervention_20210415_Pollicino_Rule_of_Law_ENG.pdf. Acesso em: 27/10/2021.

⁴⁰⁴ Tradução livre do original: “I use the term ‘regulation’ rather than the term ‘governance’ popularised in the age of networks [...]”. (ANDREWS, Leighton. Facebook ‘Regulation’: a process not a text. **UK Copyright Centre and Creative Economy Centre** (CREATe). 2020. Disponível em: <https://orca.cardiff.ac.uk/id/eprint/133226/1/CREATe%20Text.pdf>. Acesso em: 03/11/2021. p. 03).

⁴⁰⁵ ANDREWS, Leighton. Facebook ‘Regulation’: a process not a text. **UK Copyright Centre and Creative Economy Centre** (CREATe). 2020. Disponível em: <https://orca.cardiff.ac.uk/id/eprint/133226/1/CREATe%20Text.pdf>. Acesso em: 03/11/2021. p. 03.

⁴⁰⁶ ARUN, Chinmayi. Facebook’s Faces. **Harvard Law Review Forum**. v. 135, p. 236-264, 2022. Disponível em: <https://harvardlawreview.org/wp-content/uploads/2022/03/135-Harv.-L.-Rev.-F.-236.pdf>. Acesso em: 30/06/2022. p. 237.

⁴⁰⁷ Tradução livre do original: “[...] not only from its common usage in political discourse, where there is a belief that these are to a degree different things, but also from a desire to re-establish the role of the state in promoting countervailing power to achieve public value [...]” (ANDREWS, Leighton. Facebook ‘Regulation’: a process not a text. **UK Copyright Centre and Creative Economy Centre** (CREATe). 2020. Disponível em: <https://orca.cardiff.ac.uk/id/eprint/133226/1/CREATe%20Text.pdf>. Acesso em: 03/11/2021. p. 03).

⁴⁰⁸ Adotamos o termo “remédios”, cunhado por Eric Goldman – “*what I call “remedies”* –, para designar as opções que as plataformas e serviços de internet, possuem para moderar conteúdo ou contas que violam suas

solução binária – manter ou remover – como forma de fornecer opções corretivas no exercício de compensações complexas da moderação de temáticas, as quais violem as regras internas das plataformas. Portanto, os remédios apresentados são aqueles aplicados após-violação, não sendo analisado por esses dados as técnicas preventivas às normas da empresa; é preciso salientar a importância da pesquisa, mas primar pelas consequências derivadas do descumprimento às orientações para a utilização do ambiente virtual.

Destarte, verificou-se a conduta do *Facebook*, diante das postagens de Donald Trump, como resultado da desobediência às regras aplicáveis. Ato contínuo, o remédio aplicado foi a remoção do conteúdo devido à gravidade da infração, bem como a suspensão da conta por 24 horas, e, após esse período, por prazo indeterminado.⁴⁰⁹

Preliminarmente, iremos centrar a nossa construção lógica partindo do amplo debate envolto na moderação de conteúdo; analisaremos a relação binária de remoção -ou não-, do controle de conteúdo, o qual é adotado como solução primordial na maior parte dos casos ; bem como um panorama taxonômico de remédios formulados por Eric Goldmann,⁴¹⁰ mais à frente, analisaremos os fatores de escolha entre os remédios apresentados para a orientação à tomada de decisão. Por fim, haverá um levantamento normativo da moderação de conteúdo no Brasil.

3.1. Moderação de conteúdo: principais tópicos

A *web* aberta foi um planejamento almejado pelas gerações, as quais desenvolveram meios para ampliar suas aplicações – novas oportunidades de conhecimento e sociabilidade – como *home pages*, blogs pessoais, sites de notícias, *chats*, sites corporativos, jogos *online*, sites para entretenimento adulto – citando-se o conteúdo pornográfico. Essa abertura da *web*, a princípio, enfatizava o acesso ao público sem interferência de editoras e emissoras; porém, atualmente, o poder da fala seria distribuído amplamente. Contudo, com o tempo, todo esse

normas. (GOLDMAN, Eric. Content Moderation Remedies. *Michigan Technology Law Review*. v. 28, n. 1. p. 1-59, 2021. Disponível em: <https://repository.law.umich.edu/mlr/vol28/iss1/2>. Acesso em: 02/11/2021. p. 1).

⁴⁰⁹ Contudo, o Conselho de Supervisão somente confirmou a decisão de suspensão de 24 horas, quanto a suspensão por prazo indefinido, determinou que no prazo de 6 meses, o Facebook reexaminasse a penalidade, por entender inapropriada e desproporcional. Conferir na íntegra o *Case Decision 2021-001-FB-FBR*, OVERSIGHT BD. (UNITED STATES. **Facebook Oversight Board**. Case Decision 2021-001-FB-FBR, OVERSIGHT BD. 9–10. May 5, 2021. Disponível em: <https://www.oversightboard.com/sr/decision/2021/001/pdf-english>. Acesso em: 01/07/2022.

⁴¹⁰ Nesse ponto, esse capítulo irá revisar os principais pontos significativos da inovadora pesquisa de Eric Goldman - Content Moderation Remedies. *Michigan Technology Law Review*. v. 28, n. 1. p. 1-59, 2021.

universo de pessoas e informações, antes espalhadas na *web*, desaguou numa estrutura pequena composta por algumas empresas e poucas plataformas de mídia social.⁴¹¹

Essas empresas *online* desenvolveram um novo formato para manter os usuários navegando, postando, compartilhando, e demonstrando suas preferências, propensões, intenções, coletando as informações e transformando em dados personalizados, os quais foram encaixotados e comercializados para anunciantes e empresas de conteúdo. Gillespie resume: “para ficarmos livres de intermediários, aceitamos novos intermediários”;⁴¹² sendo assim, nos “livramos” da mediação da mídia “tradicional” – editoras, emissoras, jornais etc. – e transportamos esse controle para as redes sociais – *Facebook*, *Twitter*, *YouTube* etc.

Essa nova realidade trouxe outra perspectiva aos usuários em razão da possibilidade de uma “aldeia global” – metáfora de Marshall McLuhan,⁴¹³ permitindo aos indivíduos a união por interesses comuns, não apenas geograficamente, com construções sociais diversificadas e progressistas, tudo do zero. Entretanto, as consequências da globalização virtual começaram a ser vistas de forma danosa, tornando-se imperativa a instrução para um molde de governança capaz de proteger os valores das plataformas e dos usuários.⁴¹⁴

Imergindo nas atribuições desse governo virtual, vê-se a moderação como aspecto essencial à sobrevivência dessas plataformas; se, por um lado, a *Internet* facilitou a interação e a sociabilidade; por outro, tornou mais fácil o acesso para maus atores, a exemplo de *trolls*, *spammers*, hackers mal-intencionados, os quais podem impedir ou frustrar esses canais *online*. Como sinalizado por Langvardt, “imagine seu e-mail sem filtragem de spam ou seu *feed* do *Facebook* se ele fosse preenchido diariamente com vídeos de decapitação e pornografia violenta”,⁴¹⁵ seria impossível de utilizar, portanto, o trabalho da moderação é evitar que isso ocorra.

⁴¹¹ GILLESPIE, Tarleton. **Custodians of the inernet**: Platforms, content moderation, and the hiddes decisions that shape social media. Yale University Press/New Haven & London. 2018. Disponível em: https://www.researchgate.net/publication/327186182_Custodians_of_the_inernet_Platforms_content_moderation_and_the_hidden_decisions_that_shape_social_media. Acesso em: 23/11/2021. p. 15-16.

⁴¹² Tradução livre do original: “[...] to be free of intermediaries, we accepted new intermediaries.” (GILLESPIE, Tarleton. **Custodians of the inernet**: Platforms, content moderation, and the hiddes decisions that shape social media. Yale University Press/New Haven & London. 2018. Disponível em: https://www.researchgate.net/publication/327186182_Custodians_of_the_inernet_Platforms_content_moderation_and_the_hidden_decisions_that_shape_social_media. Acesso em: 23/11/2021. p. 16).

⁴¹³ MCLUHAN, Marshall. Interview with Marshall McLuhan. Entrevista concedida a L’Express. **Revista da Associação Nacional dos Programas de Pós-Graduação em Comunicação**. Brasília, v.14, n. 3, 2011. Disponível em: <https://www.e-compos.org.br/e-compos/article/download/845/610>. Acesso em: 05/07/2022.

⁴¹⁴ GILLESPIE, Tarleton. **Custodians of the inernet**: Platforms, content moderation, and the hiddes decisions that shape social media. Yale University Press/New Haven & London. 2018. Disponível em: https://www.researchgate.net/publication/327186182_Custodians_of_the_inernet_Platforms_content_moderation_and_the_hidden_decisions_that_shape_social_media. Acesso em: 23/11/2021. p. 16

⁴¹⁵ Tradução livre do original: “[...] imagine your email without spam filtering, or your Facebook feed if it were populated daily with beheading videos and violent pornography.” (LANGVARDT, Kyle. Regulating Online

Nesse raciocínio – moderação de conteúdo como uma forma de governança privada, cujo papel é ser uma “regulação” – torna-se salutar apontar o conceito desse controle, o qual, de acordo com Koop e Lodge, pode ser definido como “a intervenção intencional nas atividades de uma população-alvo, onde a intervenção é tipicamente direta – envolvendo estabelecimento de normas vinculativas, monitoramento e sanção – e exercida por atores do setor público sobre as atividades econômicas de atores do setor privado”.⁴¹⁶ De forma geral, a regulamentação é elaborada, entre outras razões, para garantir ações corretivas contra danos identificados.⁴¹⁷

Desta forma, Gillespie expõe o papel moderador é fundamental para o funcionamento efetivo das empresas de mídia social, o qual é implementado tanto pelas diretrizes adotadas, quanto pelo processo ativo de moderação de conteúdo gerenciados em relação a reclamações específicas, ou, ainda, pelo cumprimento das leis sobre conteúdo ilegal. Ademais, o maior desafio perpassa pela necessidade em definir: exatamente quando, como e por que intervir. Surgindo inevitáveis dúvidas e debates a respeito da obrigatoriedade em moldar e, às vezes, restringir o conteúdo. Essas obrigações acompanham ou contrariam a promessa legal de liberdade de expressão?⁴¹⁸

Em função do grande contingente de materiais gerados por terceiros, gerou-se a imprescindível demanda sobre as restrições das publicações (seja por meio do bloqueio ou retirada do assunto), originando a prática da moderação de conteúdo. Segundo Grimmelmann, moderação são “os mecanismos de governança que estruturam a participação em uma comunidade para facilitar a cooperação e prevenir abusos.”⁴¹⁹

Content Moderation. **The Georgetown Law Journal**. v. 106, p. 1353-1387, 2018. Disponível em: <https://www.law.georgetown.edu/georgetown-law-journal/wp-content/uploads/sites/26/2018/07/Regulating-Online-Content-Moderation.pdf>. Acesso em: 13/07/2022. p. 1359).

⁴¹⁶ Tradução livre do original: “[...] he intentional intervention in the activities of a target population, where the intervention is typically direct – involving binding standard-setting, monitoring and sanctioning – and exercised by public-sector actors on the economic activities of private-sector actors.” (KOOP, Christel; LODGE, Martin. What is regulation? An interdisciplinary concept analysis. **Regulation and Governance**. v. 11, n. 1, p. 95-108, 2017. Disponível em: http://eprints.lse.ac.uk/62135/1/Lodge_What_is_regulation_2016.pdf. ou DOI: 10.1111/rego.12094. Acesso em: 02/07/2022. p. 21)

⁴¹⁷ ANDREWS, Leighton. Facebook ‘Regulation’: a process not a text. **UK Copyright Centre and Creative Economy Centre (CREATe)**. 2020. Disponível em: <https://orca.cardiff.ac.uk/id/eprint/133226/1/CREATe%20Text.pdf>. Acesso em: 03/11/2021. p. 14.

⁴¹⁸ GILLESPIE, Tarleton. **Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media**. Yale University Press/New Haven & London. 2018. Disponível em: https://www.researchgate.net/publication/327186182_Custodians_of_the_internet_Platforms_content_moderation_and_the_hidden_decisions_that_shape_social_media. Acesso em: 23/11/2021. p. 5 e seguintes.

⁴¹⁹ Tradução livre do original: “[...] the governance mechanisms that structure participation in a community to facilitate cooperation and prevent abuse.” (Grimmelmann, James. The Virtues of Moderation, **The Yale Journal of Law & Technology**. v. 17, n. 1, p. 42-109, 2015. p. 47. Disponível em: <https://core.ac.uk/download/pdf/72836082.pdf>. ou <https://digitalcommons.law.yale.edu/yjolt/vol17/iss1/2>. Acesso em: 10/01/2022.).

Nesse sentido, Eric Goldman sustenta, devido às consequências graves de contornos de difícil previsão, o senso comum como desenvolvedor do critério, quase automático, de: um conteúdo ou conta que violem regras – seja legal ou ilegal – deve ser imediatamente removido da plataforma⁴²⁰. Diante disso, percebe-se essa “remoção” como aplicação de um remédio corretivo, numa abordagem “binária”, observando-se, por esse prisma, a decisão evidenciada por analogia a um, botão: liga/desliga.

Nota-se a pressão sob os reguladores e plataformas ocasionando um receio em verificar outras possibilidades de decisão na correção de uma temática oposta às normas definidas. Contudo, as plataformas têm desenvolvido e experimentando outras soluções além da remoção. Exemplo disso, relembra-se o caso “*Tramp Ban*”, no qual houve a aplicação de remédios corretivos alternativos, conforme é possível constatar da lista de ações empregadas por diversas plataformas *online* no aludido caso. Abaixo cita-se um recorte da última atualização, apresentadas na ordem decrescente:⁴²¹

Platform actions in response to January 6 Capitol events - Newest to Oldest					
The following chart is meant to help document how social media platforms have responded to the January 6 events at the US Capitol. Where possible, we have included direct statements from the platforms or their spokespeople, links to specific policies they have referenced, and archives of any content that was affected by platform blocks or takedowns. Please see the “ Column Key ” below for a description of each.					
Platform / Date	Action	Link to statement	Policy referenced	What triggered the action?	Archived links (if applicable)
 YouTube 5:21 pm January 19, 2021	Ban Donald J Trump channel from uploading content for at least seven more days	Confirmed to a statement to Reuters	Concerns over potential for violence		
 Facebook 8:00 am, January 10, 2021	Expanded ban on weapons and weapon accessories to include products “gun safes, vests and gun holders”	Updated statement on preparations for inauguration	No specific policy mentioned. Facebook already prohibits ads for weapons, ammunition and accessories like silencers		

Figura 2: Ações adotadas pelas plataformas sobre o evento ocorrido no Capitólio Americano. Fonte: https://docs.google.com/document/d/1dNC87RtdPWBXXReTsrAl-Sknw4PtwanPX0CA_oi20ec/edit?fbclid=IwAR05m4XHSS-H2znFuSKlIGhBN0FnQqUoqgxau0vbQOST-yEMF9aCnnoPM9w#.

Compreende-se a moderação de conteúdo como objeto de longos e contínuos debates, não havendo, ainda, consensos sobre suas abordagens. Com isso, apresentamos os principais tópicos dessas discussões sobre o controle de temáticas postadas, além dos já analisados nesse estudo, como a governança de plataforma, explicitando o não aprofundamento direto de

⁴²⁰ GOLDMAN, Eric. Content Moderation Remedies. *Michigan Technology Law Review*. v. 28, n. 1. p. 1-59, 2021. p. 5. Disponível em: <https://repository.law.umich.edu/mtlr/vol28/iss1/2>. Acesso em: 02/11/2021.

⁴²¹ Consulte a lista completa em: https://docs.google.com/document/d/1dNC87RtdPWBXXReTsrAl-Sknw4PtwanPX0CA_oi20ec/edit?fbclid=IwAR05m4XHSS-H2znFuSKlIGhBN0FnQqUoqgxau0vbQOST-yEMF9aCnnoPM9w#.

outros métodos nesta pesquisa, tais como responsabilidade algorítmica,⁴²² legitimidade,⁴²³ atuação do Estado,⁴²⁴ neutralidade da rede⁴²⁵, dentre outros, os quais certamente, integram toda essa robusta discussão.

O cerne passa a ser o apresentado por Eric Goldman, acrescentando-se um quarto dentre aqueles abordados, de forma geral, pela doutrina. Contextualizando: Que conteúdo e atividade devem ser permitidos *online*? Quem deve fazer as regras substantivas de conteúdo e atividades *online*? Quem deve determinar se ocorreu uma violação de regra e quem deve ouvir quaisquer apelações dessas decisões?⁴²⁶ Como regular a liberdade de expressão na era digital?

a. *Que conteúdo e atividade devem ser permitidos online?*

Essa questão envolve aspectos substantivos de regras de moderação, como o conteúdo e atividades. Inclusive, a alteração da atuação de padrões para fixação de regras representa um marco na evolução das redes sociais; de acordo com Klonick, foi essa mudança de paradigma na governança orientada por *standards* normativos o impulso para um novo sistema baseado em regras:

A moderação de conteúdo no YouTube e no Facebook evoluiu de um sistema inicial de padrões para um intrincado sistema de regras devido (1) ao rápido aumento de usuários e volume de conteúdo; (2) a globalização e a diversidade da comunidade online; e (3) a crescente dependência de equipes de moderadores humanos com diversas formações.⁴²⁷

⁴²² Ver COBBE, Jennifer. Algorithmic Censorship by Social Platform: Power and Resistance. **Philosophy & Technology**. v. 34, p. 739-766, 2021. Disponível em: <https://doi.org/10.1007/s13347-020-00429-0>. Acesso em: 26/08/2021; e MICKLITZ, Hans-W.; VILLANUEVA, Aurélie Anne. Responsibilites of Companies in the Algorithmic Society. In. MICKLITZ, Hans-W.; POLLICINO, Oreste, REICHMAN, Amnon, SIMONCINI, Andrea, SARTOR, Giovanni; DE GREGORIO, Gregorio. (Eds.). **Constitutional Challenges in the Algorithmic Society**. Cambridge: Cambridge University Press, 2021. Disponível em: <https://doi.org/10.1017/9781108914857>. Acesso em: 23/11/2021.

⁴²³ Conferir: JAMART, Anne-Cleire. **Legitimacy in internet governance: Multistakeholderism and global constituent Power**. Dissertação (Mestrado em Direito). Ithaca, NY: Faculty of the Graduate School of Cornell University, 2021; e SUZOR, Nicolas. Digital Constitutionalism: Using the Rule to Evaluate the Legitimacy of Governance by Platforms. **Social Media + Society**. v. 4, N. 3, July 2018. Disponível em: <https://journals.sagepub.com/doi/pdf/10.1177/2056305118787812>. Acesso em: 18/02/2021.

⁴²⁴ Ver: LOO, Rory Van. The Missing Regulatory State: Monitoring Business in an Age of Surveillance. **Vanderbilt Law Review**. v. 72, n. 5, p. 1563-1631, 2019. Disponível em: https://scholarship.law.bu.edu/faculty_scholarship/678. Acesso em: 03/11/2021; e EILKIN-KOREN, Niva; PEREL, Maayan. Guarding the Guardians: content moderation by online intermediaries and the Rule of Law. In: FROSIO, Giancarlo (Ed.). **The Oxford Handbook of Online Intermediary Liability**. New York, NY: Oxford University Press, 2020. Disponível em: <https://ssrn.com/abstract=3542992>. Acesso em: 28/10/2021.

⁴²⁵ PINHEIRO, Guilherme Pereira. **Liberdade de Expressão e Neutralidade de Rede na Internet**. Rio de Janeiro: Lumen Juris, 2017. e SOUZA, Carlos Affonso Pereira de.; MONIZ, Pedro de Paranaguá; VIEIRA JÚNIOR, Sérgio Branco. Neutralidade da Rede, Filtragem de Conteúdo e Interesse Público: reflexões sobre o bloqueio do site Youtube no Brasil. **Revista de Direito Administrativo - RDA**, ano 16, n. 246, p. 50-78, 2007. Disponível em: <https://www.forumconhecimento.com.br/periodico/125/191/2310>. Acesso em: 08/07/2022.

⁴²⁶ GOLDMAN, Eric. Content Moderation Remedies. **Michigan Technology Law Review**. v. 28, n. 1, p. 1-59, 2021. p. 7-8. Disponível em: <https://repository.law.umich.edu/mltr/vol28/iss1/2>. Acesso em: 02/11/2021.

⁴²⁷ Tradução livre do original: “Content moderation at YouTube and Facebook developed from an early system of standards to an intricate system of rules due to (1) the rapid increase in both users and volume of content; (2)

Nesse ritmo, foram desenvolvidas regras proibitivas de pornografia infantil, discurso de ódio, violação de direitos autorais, desinformação, exploração sexual, suicídio, exploração humana, dentre outros.

O *Facebook* informou em seu site, que, somente nos primeiros meses de 2022 – de janeiro a março – somam-se cerca de 15,1 milhões de conteúdos com medidas punitivas tomadas, somente em relação ao discurso de ódio. O número de itens de conteúdo (como publicações, fotos, vídeos ou comentários) foi contabilizado ou as contas cujas ações violassem padrões foram suspensas, bloqueadas, ou ainda banidas. Ademais, o termo “Tomar medidas” pode incluir remover um item do conteúdo do *Facebook* ou *Instagram*, cobrir com um aviso fotos ou vídeos com conteúdos perturbadores para alguns públicos ou desativar contas. Quando o conteúdo é encaminhado para alguma autoridade para tomar medidas judiciais ou extrajudiciais cabíveis, ele não é contabilizado.⁴²⁸ Registre-se que o tema, no qual se debatem o conteúdo e as atividades, possui um amplo e complexo rol de itens os quais pudessem ser permitidos no ambiente *online*.

b. Quem deve fazer as regras substantivas de conteúdo e atividades online?

Outro ponto bastante analisado pela doutrina trata da formulação de regras cuja função é primordialmente do Estado, concebendo-se por meio de um ato legiferante formal ou por determinações judiciais – ao interpretarem leis ou conformarem-se a decisões as quais firmaram jurisprudência àqueles tipos de conteúdo e atividades. Por outro lado, embora possam ser ilegais, alguns conteúdos podem ter o seu teor punido pela própria plataforma seja a sua natureza menos incisiva a fim de tomar medidas mais céleres para corrigir de caráter tormentoso ou terrível.⁴²⁹

Destarte, o caráter resolutivo das empresas de mídias sociais pode, por si, estabelecer suas próprias regras substantivas, sejam quanto ao conteúdo e quanto às atividades de seus serviços.⁴³⁰ Contudo, os Termos de Serviço e/ou Uso devem complementar as regras oriundas do Estado.

the globalization and diversity of the online community; and (3) the increased reliance on teams of human moderators with diverse backgrounds.” (KLONICK, Kate. *The New Governors: The People, Rules and Processes Governing Online Speech*. **Harvard Law Review**. v. 131. p. 1598-1670. 2018. p. 1635. Disponível em: https://harvardlawreview.org/wp-content/uploads/2018/04/1598-1670_Online.pdf. Acesso em: 28/05/2022).

⁴²⁸ META. Hate Speech. Community Standards Enforcement Report. 2022. Disponível em: <https://transparency.fb.com/data/community-standards-enforcement/hate-speech/facebook/>. Acesso em: 08/07/2022.

⁴²⁹ GOLDMAN, Eric. Content Moderation Remedies. **Michigan Technology Law Review**. v. 28, n. 1. p. 1-59, 2021. Disponível em: <https://repository.law.umich.edu/mtlr/vol28/iss1/2>. Acesso em: 02/11/2021. p. 7-8.

⁴³⁰ Tratamos sobre o tema no Capítulo II, item 2.4.

c. *Quem deve determinar se ocorreu uma violação de regra e quem deve ouvir quaisquer apelações dessas decisões?*

Nesse quesito, como explica Goldman, historicamente quem estipula a violação das regras ou leis são os Tribunais, quando acionados por algum usuário, ou mesmo, por outras entidades públicas, decidindo sobre tais questões de acordo com o arcabouço legal aplicável ao caso ou, ainda, se o resultado da situação em debate possa ser de custo elevado para as empresas. Contudo, dada a necessidade de celeridade e de atender a demandas dos usuários a contento, as plataformas de redes sociais têm desenvolvido tanto suas regras com maior efetividade e abrangência, como tomado as decisões sobre o conteúdo ou atividade *online*, se violaram ou não as regras da empresa, regras de cunho social ou de apelo à moralidade pública, inclusive, algumas dessas decisões possuem uma instância de revisão, como,⁴³¹ por exemplo, o *Facebook Oversight Board* – FOB.

O *Conselho de Supervisão do Facebook*, como também é conhecido o FOB originou-se a partir de uma declaração, em 2018, de Zuckerberg, sobre uma estrutura semelhante ao “Supremo Tribunal”, como resposta às intensas críticas exigindo maior responsabilidade do usuário como também o usuário exigindo maior responsabilidade empresarial e social pela plataforma. Com isso, anunciou em novembro de 2018 o *Facebook* anunciou a criação do comitê de “Governança e Supervisão Independentes” cuja instalação e funcionamento ocorreriam até o final de 2019.⁴³²

Após, o pronunciamento, a empresa iniciou a fase de consulta global e também o recrutamento. Com os resultados da consultoria global, foram publicados, num relatório executivo em 27 de junho de 2019, os itens de conformidade pelos usuários, embora não se chegado a um consenso, o relatório serviu de embasamento e de estudo de tendência do conjunto de valores substantivos advindos dos seguidores e de sustentáculo ao propósito do Conselho, o qual deve se aprimorar, pois, muitos dos, valores são dinâmicos e serão definidos ou rejeitados de acordo com a latência do próprio ambiente da plataforma.

Passada essa etapa, iniciou-se a fase de estrutura e projeto, anunciando-se a minuta de um Estatuto de convivência na plataforma, seguindo-se à implementação. Nos estágios finais da redação e implementação da Carta, foi inaugurado um Conselho de Confiança e lançado o Contrato de Responsabilidade Limitada do Conselho de Supervisão, formando uma *Limited*

⁴³¹ GOLDMAN, Eric. Content Moderation Remedies. **Michigan Technology Law Review**. v. 28, n. 1. p. 1-59, 2021. Disponível em: <https://repository.law.umich.edu/mtlr/vol28/iss1/2>. Acesso em: 02/11/2021. p. 9.

⁴³² KLONICK, Kate. The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression. **The Yale Law Journal**. v. 129, p. 2418-2499, 2020. Disponível em: https://scholarship.law.stjohns.edu/cgi/viewcontent.cgi?article=1312&context=faculty_publications. Acesso em: 20/06/2022. p. 2425.

Liability Company (LLC), sociedade de responsabilidade limitada. Como forma de conceder autonomia e independência ao conselho.⁴³³

Todos os documentos criados foram trazidos ao público da plataforma em 28 de janeiro de 2020, e o Estatuto foi o documento que fundou, finalmente, o Conselho. Estão divididos em quatro artigos esmiuçando a relação do Conselho com o *Facebook*, o *Trust*, o Conselho de Supervisão e as Pessoas que poderão recorrer.⁴³⁴ Após a operação iniciada ainda em 2020, os primeiros casos de impropriedade diante das regras do Estatuto foram inicialmente analisados em outubro do mesmo ano.⁴³⁵

É inevitável não fazer analogia entre o *Facebook Oversight Board* e a Suprema Corte,⁴³⁶ decorrendo daí o “Supremo Tribunal do *Facebook*”, termo utilizado por Sofia Ranchordàs, que afirma ser o FOB, a espécie de um “Chefe de justiça” da plataforma, pois, em sua visão, o *Facebook* atua como um “Estado *online*” no qual “decide o recurso e tem o poder de aceitar ou recusar casos, e é composto em parte por membros independentes, e por advogados, jornalistas e ativistas de internacionalmente renomados”.⁴³⁷ Complementando a

⁴³³ KLONICK, Kate. The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression. **The Yale Law Journal**. v. 129, p. 2418-2499, 2020. Disponível em: https://scholarship.law.stjohns.edu/cgi/viewcontent.cgi?article=1312&context=faculty_publications. Acesso em: 20/06/2022. p. 2469.

⁴³⁴ KLONICK, Kate. The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression. **The Yale Law Journal**. v. 129, p. 2418-2499, 2020. Disponível em: https://scholarship.law.stjohns.edu/cgi/viewcontent.cgi?article=1312&context=faculty_publications. Acesso em: 20/06/2022. p. 2469.

⁴³⁵ META. Criação do Conselho de Supervisão. 2022. Disponível em: <https://transparency.fb.com/pt-br/oversight/creation-of-oversight-board/>. Acesso em: 08/07/2022.

⁴³⁶ Sobre esse tema conferir as colocações de Kate Klonick: “A analogia com os tribunais é valiosa, mas também imperfeita. Comparado aos tribunais federais dos EUA, a capacidade do Conselho de selecionar casos e seu poder jurisdicional de autoridade final são semelhantes aos da Suprema Corte dos EUA. O uso de painéis de cinco pessoas que decidem por maioria é semelhante a um tribunal de circuito, embora a aprovação en banc do Conselho de uma decisão não seja exigida por um tribunal de circuito. Como um tribunal de primeira instância, o Conselho – principalmente por meio de sua equipe administrativa – também pesquisa e constrói casos em torno de fatos. E, assim como nos tribunais, as partes – o usuário apelante e o próprio Facebook – apresentam argumentos por escrito. O escopo jurisdicional limitado do Conselho o torna semelhante a um tribunal de jurisdição especial ou exclusiva.” Tradução livre do original: “The analogy to courts is valuable, but also imperfect. Compared to U.S. federal courts, the Board’s ability to select cases and its jurisdictional power of final authority are similar to that of the U.S. Supreme Court. The use of five-person panels that decide by majority is akin to a circuit court, though the Board’s en banc approval of a decision is not required by a circuit court. Like a trial court, the Board—primarily through its administrative staff—also researches and builds cases around facts. And as with courts, parties—the appealing user and Facebook itself—submit written arguments. The Board’s limited jurisdictional scope renders it similar to a court of special or exclusive jurisdiction.” (KLONICK, Kate. The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression. **The Yale Law Journal**. v. 129, p. 2418-2499, 2020. Disponível em: https://scholarship.law.stjohns.edu/cgi/viewcontent.cgi?article=1312&context=faculty_publications. Acesso em: 20/06/2022. p. 2476)

⁴³⁷ Tradução livre do original: “[...] beslist in hoger beroep, heeft de bevoegdheid om zaken te accepteren of te weigeren en bestaat voor een deel uit onafhankelijke leden. De leden van de FOB zijn internationale gerenommeerde juristen, journalisten en activisten.” (RANCHORDAS, Sofia, *Digitale platformen: Globale staten naast de staat? Digital Platforms: Global States beyond the State?* **University of Groningen Faculty of Law** Research Paper N. 35/2020, In Besselink, L. (red) *De economische constitutie*. Staatsrechtsconferentie

analogia afirmando que, com a instituição desse “tribunal”, o *Facebook* está introduzindo um sistema de freios e contrapesos no “*Facebookistan*” – expressão utilizada por Evelyn Douek para exprimir o domínio soberano do *Facebook*.⁴³⁸

Nesse mesmo estudo, Douek expõe o verdadeiro valor do *Facebook Oversight Board* ao instituir, para o ecossistema de moderação de conteúdo do *Facebook* uma balança atuante entre os dois extremos: de um lado a correção de erros individuais, e de outro o estabelecimento de regras de fala globalmente aplicável.⁴³⁹ Nesse desiderato, o *Facebook Oversight Board*, contribui para construção de duas funções simples, mas de suma importância:

Primeiro, pode ajudar a destacar pontos fracos no processo de formação de políticas no Facebook, removendo bloqueios (como pontos cegos e inércia) no “processo legislativo” que leva à formulação de seus Padrões da Comunidade. Em segundo lugar, ao fornecer um fórum independente para a discussão de decisões de moderação de conteúdo contestado, o FOB pode ser um fórum importante para o processo de raciocínio público necessário para que as pessoas de uma comunidade pluralista aceitem as regras que as regem, mesmo que não concordem com a substância dessas regras.⁴⁴⁰

Conforme consta no sítio da *Meta*, o Conselho de Supervisão foi criado com a função de oferecer verificação independente sobre algumas decisões, dotadas de maior relevância a qual compreenda grau de complexidade mais elevada na tomada da decisão sobre o conteúdo. Afirmar ainda que a ideia inicial do Conselho é bastante simples: “Não cabe ao *Facebook* tomar, por conta própria, tantas decisões importantes sobre a liberdade de expressão e a segurança.”⁴⁴¹

Segundo Pollicino e De Gregorio, o *Facebook Oversight Board* (FOB) representa aos usuários um novo remédio potencial contra a remoção discricionária e arbitrária de conteúdo,

2019, University of Amsterdam Press, 2020. Disponível em: <https://ssrn.com/abstract=3663065>. Acesso em: 28/10/2021. p. 2)

⁴³⁸ DOUEK, Evelyn. Facebook’s oversight board: move fast with stable infrastructure and humility’. **North Carolina Journal of Law & Technology**. v. 21, n. 1, p. 1-79, 2019. Disponível em: <https://scholarship.law.unc.edu/cgi/viewcontent.cgi?article=1389&context=ncjolt>. Acesso em: 03/11/2021. p. 3.

⁴³⁹ DOUEK, Evelyn. Facebook’s oversight board: move fast with stable infrastructure and humility’. **North Carolina Journal of Law & Technology**. v. 21, n. 1, p. 1-79, 2019. Disponível em: <https://scholarship.law.unc.edu/cgi/viewcontent.cgi?article=1389&context=ncjolt>. Acesso em: 03/11/2021. p. 7;

⁴⁴⁰ Tradução livre do original: “First, it can help highlight weaknesses in the policy formation process at Facebook, removing blockages (such as blind spots and inertia) in the “legislative process” leading to the formulation of its Community Standards. Second, by providing an independent forum for the discussion of disputed content moderation decisions, the FOB can be an important forum for the process of public reasoning necessary for persons in a pluralistic community to come to accept the rules that govern them, even if they disagree with the substance of those rules.” (DOUEK, Evelyn. Facebook’s oversight board: move fast with stable infrastructure and humility’. **North Carolina Journal of Law & Technology**. v. 21, n. 1, p. 1-79, 2019. Disponível em: <https://scholarship.law.unc.edu/cgi/viewcontent.cgi?article=1389&context=ncjolt>. Acesso em: 03/11/2021. p. 7).

⁴⁴¹ META. Criação do Conselho de Supervisão. 2022. Disponível em: <https://transparency.fb.com/pt-br/oversight/creation-of-oversight-board/>. Acesso em: 08/07/2022.

pois, tornando a tomada de decisão do *Facebook*, sobre moderação de conteúdo mais responsável e transparente. Além disso, enseja uma nova era de adjudicação privada de conteúdo *online* e, em sua visão, será um sistema alternativo para fazer cumprir os direitos humanos em escala global,⁴⁴² acrescentando: “constitucionalismo digital oferece uma estrutura para olhar para esta nova forma de adjudicação privada de conteúdo *online* e seus desafios”.⁴⁴³

Nesse sentido, constatou-se uma deliberação concreta, e não uma mera abstração, conforme foi percebido a partir dos temas relevantes já enfrentados, como, por exemplo: o genocídio em Mianmar, os recentes eventos nos Estados Unidos e a pandemia destacaram como “o ordenamento privado pelas redes sociais tem impacto sobre os direitos individuais e a democracia.”⁴⁴⁴

Para Balkin, o Conselho de Supervisão para decisões de conteúdo do *Facebook* representa uma estratégia para gerar confiança pública, com isso, busca estabelecer um tipo de legitimidade para suas decisões de moderação de conteúdo, inclusive, afirmando o seguinte: “não tenho nenhuma objeção ao Conselho em teoria. Devemos encorajar toda reforma que dê às empresas de mídia social incentivos para agir em público em relação à moda.”⁴⁴⁵

Contudo, Balkin, endereça críticas ao *Facebook* no seguinte vive num “ciclo de envolvimento em mau comportamento, ser pego, e desculpar-se profusamente e prometer consertar seus caminhos.”⁴⁴⁶ Isso não por ser incompetente, mas por causa do “desalinhamento fundamental de incentivos entre seus objetivos e as necessidades do público e porque tem um conflito de interesses inerente com seus usuários finais e, na verdade, com a própria democracia.”⁴⁴⁷

⁴⁴² POLLICINO, Oreste; DE GREGORIO, Giovanni: Shedding Light on the Darkness of Content Moderation: The First Decisions of the Facebook Oversight Board, **VerfBlog**, 2021/2/05, disponível em: <https://verfassungsblog.de/fob-constitutionalism/>, DOI: 10.17176/20210205-235959-0. Acesso em 24/11/2021.

⁴⁴³ POLLICINO, Oreste; DE GREGORIO, Giovanni: Shedding Light on the Darkness of Content Moderation: The First Decisions of the Facebook Oversight Board, **VerfBlog**, 2021/2/05, disponível em: <https://verfassungsblog.de/fob-constitutionalism/>, DOI: 10.17176/20210205-235959-0. Acesso em 24/11/2021.

⁴⁴⁴ POLLICINO, Oreste; DE GREGORIO, Giovanni: Shedding Light on the Darkness of Content Moderation: The First Decisions of the Facebook Oversight Board, **VerfBlog**, 2021/2/05, disponível em: <https://verfassungsblog.de/fob-constitutionalism/>, DOI: 10.17176/20210205-235959-0. Acessado em 24/11/2021.

⁴⁴⁵ Tradução livre do original: “I have no objection to the Board in theory. We should encourage every reform that gives social media companies incentives to act in a public regarding fashion”. (BALKIN, Jack M., How to Regulate (and Not Regulate) Social Media. **Journal of Free Speech Law**. v. 1, p.71-96, Knight Institute Occasional Paper Series, No. 1, Yale Law School, Public Law Research Paper. 2021. Disponível em: <https://ssrn.com/abstract=3484114>. Acesso em: 10/12/2021. p. 88)

⁴⁴⁶ Tradução livre do original: “[...] cycle of engaging in bad behavior, getting caught, apologizing profusely and promising to mend its ways”. (BALKIN, Jack M., How to Regulate (and Not Regulate) Social Media. **Journal of Free Speech Law**. v. 1, p.71-96, Knight Institute Occasional Paper Series, No. 1, Yale Law School, Public Law Research Paper. 2021. Disponível em: <https://ssrn.com/abstract=3484114>. Acesso em: 10/12/2021. p. 88)

⁴⁴⁷ Tradução livre do original: “[...] fundamental misalignment of incentives between its goals and the public’s needs, and because it has an inherent conflict of interest with its end users and, indeed, with democracy itself.

Ademais, preconiza empresas com fins lucrativos como o *Facebook*, normalmente procuram externalizar o máximo possível de custos de suas atividades para outras pessoas. Seus modelos de negócios não se importam com sua democracia. Para reverter esse quadro, Balkin indica, a adoção de normas profissionais voltadas para o público, e internalizar alguns dos custos de sua moderação de conteúdo, além de outros.⁴⁴⁸

d. *Como regular a liberdade de expressão na era digital?*

Essa pergunta, por si, demandaria um estudo próprio. Esposaremos, para os propósitos desse estudo, como mais um dos pontos importantes debatidos na doutrina. Segundo Bowers e Zittrain, existem três eras de governança do espaço digital.⁴⁴⁹ Sendo a primeira, a Era dos Direitos, com início na expansão da *internet* nos anos 90, apresentando como resposta inaugural aos novos desafios à autorregulação, se levássemos em consideração a abstenção estatal, a medida necessária para a consolidação e o fortalecimento da *internet* ainda está em processo de amadurecimento.⁴⁵⁰ Essa visão, além de promover a inovação, era compatível à realidade do ambiente virtual naquele momento, no qual o conteúdo publicado na rede possuía um alcance mais restrito, centrado em blogs, fóruns, chats ou já produzidos pelas grandes mídias tradicionais, as quais foram incorporadas ambiente *online*.⁴⁵¹

A segunda era teve início a partir de 2010, sendo impulsionada pela ascensão das plataformas, as quais passaram a indexar e moderar as temáticas, imiscuindo a divisão anterior de conteúdo produzido por usuários e conteúdo da mídia tradicional.⁴⁵² Por conseguinte, o assunto passou a ser massificado e, independentemente de sua origem, era consumido de igual forma, tornando-se viral em um curto espaço de tempo, sem qualquer análise de boa-fé ou cautela de seu autor. Com isso, passou a refletir em temas

[...].” (BALKIN, Jack M., How to Regulate (and Not Regulate) Social Media. **Journal of Free Speech Law**. v. 1, p.71-96, Knight Institute Occasional Paper Series, No. 1, Yale Law School, Public Law Research Paper. 2021. p. 89. Disponível em: <https://ssrn.com/abstract=3484114>. Acesso em: 10/12/2021).

⁴⁴⁸ BALKIN, Jack M., How to Regulate (and Not Regulate) Social Media. **Journal of Free Speech Law**. v. 1, p.71-96, Knight Institute Occasional Paper Series, No. 1, Yale Law School, Public Law Research Paper. 2021. p. 89. Disponível em: <https://ssrn.com/abstract=3484114>. Acesso em: 10/12/2021.

⁴⁴⁹ BOWERS, Jonh; ZITTRAIN, Jonathan. Answering impossible questions: content governance in an age of disinformation. **Harvard Kennedy School (HKS) Misinformation Review**. v. 1, n. 1. 2020. Disponível em: <https://misinforeview.hks.harvard.edu/article/content-governance-in-an-age-of-disinformation/>. Acesso em: 08/08/2022.

⁴⁵⁰ BOWERS, Jonh; ZITTRAIN, Jonathan. Answering impossible questions: content governance in an age of disinformation. **Harvard Kennedy School (HKS) Misinformation Review**. v. 1, n. 1. 2020. Disponível em: <https://misinforeview.hks.harvard.edu/article/content-governance-in-an-age-of-disinformation/>. Acesso em: 08/08/2022.

⁴⁵¹ BARROSO, Luna van Brussel. **Liberdade de expressão e democracia na Era Digital**: o impacto das mídias sociais no mundo contemporâneo. Belo Horizonte: Fórum, 2022. p. 219

⁴⁵² BOWERS, Jonh; ZITTRAIN, Jonathan. Answering impossible questions: content governance in an age of disinformation. **Harvard Kennedy School (HKS) Misinformation Review**. v. 1, n. 1. 2020. Disponível em: <https://misinforeview.hks.harvard.edu/article/content-governance-in-an-age-of-disinformation/>. Acesso em: 08/08/2022.

sensacionalistas e difamatórios, servindo para manipular os indivíduos inseridos naquele ambiente. Desta forma, como sustentam Abboud, Nery Júnior e Campos, é perceptível, nesse enfoque, a impossibilidade de transferir ao Estado a culpabilidade sobre as ameaças a direitos fundamentais na *internet*, bem como uma postura inerte do poder estatal não é suficiente para proteger, nesse ambiente hostil, os interesses públicos relevantes, como a liberdade de expressão, a privacidade e outros⁴⁵³. Zittrain caracteriza essa “era” de saúde pública por meio dessas preocupações e como devemos combater mais se assemelham aos assuntos relacionados a cuidados salutareis; distanciando-se, desta forma, da analogia sobre um discurso de direitos, tendo em vista a implicação da violação mútua (usuário versus usuário), sendo este o parecer técnico sob a ótica da rede.⁴⁵⁴

A era do Processo é a última; seu significado é substancialmente ligado ao período atual. As peculiaridades abordadas nessa época são impulsionadas por um movimento crítico ao desempenho passivo das plataformas e do Estado. Surge, portanto, sobre a desregulação, o debate público como fator propulsor para os prejuízos individuais e institucionais.⁴⁵⁵ Zittrain explica a inexistência de ligação entre essa postura e as alterações legislativas, tendo em vista a insuficiência estatal, por si só, em garantir a efetiva e abrangente proteção aos direitos fundamentais, sobretudo à liberdade de expressão.⁴⁵⁶

Nesse contexto, Luna Barro assevera a regulação puramente estatal e autorregulação insuficientes aos fins garantistas, adicionando a ideia da *autorregulação regulada* se mostrar como uma alternativa promissora. O objetivo desse modelo é explorar a cooperação entre Estado e atores regulados, aliando-se ao conhecimento dos agentes privados, a fim de garantir os direitos fundamentais e assegurar o interesse público.⁴⁵⁷

⁴⁵³ ABBOUD, Georges; NERY JÚNIOR, Nelson; CAMPOS, Ricardo. **Fake news e regulação**. São Paulo: Thomson Reuters Brasil, 2021. p. 125.

⁴⁵⁴ ZITTRAIN, Jonathan. **Three Eras of Digital Governance**. Disponível em: <https://ssrn.com/abstract=3458435> ou <http://dx.doi.org/10.2139/ssrn.3458435>. Acesso em: 08/07/2022.

⁴⁵⁵ BOWERS, Jonh; ZITTRAIN, Jonathan. Answering impossible questions: content governance in an age of disinformation. **Harvard Kennedy School (HKS) Misinformation Review**. v. 1, n. 1. 2020. Disponível em: <https://misinforeview.hks.harvard.edu/article/content-governance-in-an-age-of-disinformation/>. Acesso em: 08/08/2022.

⁴⁵⁶ Luna Barroso, aponta os motivos, que enumeramos: i. a regulação deve passar necessariamente pelos atores intermediários; ii. a vida digital está em constante e rápida evolução, e um regulação estatal específica, ficaria desatualizada e ineficaz, rapidamente; iii. a maioria dos entes estatais não possuem conhecimento técnico específico, sobre as plataformas e algoritmos; iv. a flagrante divergência entre como estipular se determinado conteúdo é ilícito, e com deve ser removido. (BARROSO, Luna van Brussel. **Liberdade de expressão e democracia na Era Digital**: o impacto das mídias sociais no mundo contemporâneo. Belo Horizonte: Fórum, 2022. p. 219-220).

⁴⁵⁷ BARROSO, Luna van Brussel. **Liberdade de expressão e democracia na Era Digital**: o impacto das mídias sociais no mundo contemporâneo. Belo Horizonte: Fórum, 2022. p. 223)

Entretanto, a busca por um modelo perfeito de regulação converte-se em aspirações utópicas, levando em consideração a inevitabilidade da existência de pontos cegos ao controle, ações discricionárias de censura a conteúdos lícitos e legítimos, bem como a insuficiência de restrições, as quais culminariam na visibilidade de conteúdos ilícitos.

Balkin, verbera o envolvimento imprescindível das plataformas de mídia social na moderação de conteúdo, as quais devem exercer o controle sob todas as hipóteses, organizado ou não. Disserta, também: “os governos devem respeitar o papel das mídias sociais como curadores e editores do discurso público. Respeitar esse papel significa que a mídia social deve ter direitos editoriais, que são um subconjunto dos direitos de liberdade de expressão.”⁴⁵⁸

Destarte, observa-se a inviabilidade da irrepreensível gestão sobre as temáticas abordadas em ambientes virtuais – pronta e acabada – referentes à liberdade de expressão, porque sempre haverá um desacordo razoável em determinados assuntos, como, por exemplo, a licitude ou ilicitude de um discurso específico.

Devemos, portanto, primar por um modelo de regulação capaz de otimizar a ponderação entre direitos fundamentais de usuários e plataformas, reconhecendo a existência de possíveis de erro ou assunto, sobre os quais o consenso não será possível, como afirma Douek:

A moderação de conteúdo é inerentemente sistêmica. Os sistemas de moderação de conteúdo não prometem acertar todas as decisões individuais de fala; eles são projetados para aumentar a probabilidade de que a maioria das decisões esteja certa na maioria das vezes e, quando o sistema erra, ele o faz em uma direção preferencial. Uma abordagem sistêmica aceita a inevitabilidade dos erros e os considera no projeto de governança.⁴⁵⁹

Portanto, como explica Luna Barroso, o foco da regulação deve ser no desenvolvimento e aperfeiçoamento de um procedimento adequado e seguro, o qual proporcione a tomada de decisões sobre moderação de conteúdo, sendo suficientemente capaz de minimizar erros e legitimar essas decisões, mesmo não havendo concordância plena de seus resultados, configurando-se numa proposta de regulação em três pontos:

⁴⁵⁸ Tradução livre do original: “Accordingly, governments should respect social media’s role as curators and editors of public discourse. Respecting that role means that social media should have editorial rights, which are a subset of free speech rights.” (BALKIN, Jack M., How to Regulate (and Not Regulate) Social Media. **Journal of Free Speech Law**. v. 1, p. 71-96, Knight Institute Occasional Paper Series, No. 1, Yale Law School, Public Law Research Paper. 2021. p. 90. Disponível em: <https://ssrn.com/abstract=3484114>. Acesso em: 10/12/2021).

⁴⁵⁹ Tradução livre do original: “Content moderation is inherently systemic. Content moderation systems do not promise to get every individual speech decision right; they are designed to increase the probability that most decisions will be right most of the time and when the system errs, it does so in a preferred direction. A systemic approach accepts the inevitability of errors and factors them into governance design.” (DOUEK, Evelyn. Governing online speech: from “posts-as-trumps” to proportionality and probability. **Columbia Law Review**. v. 121, n. 3, p. 759-837, 2021. p. 791. Disponível em: https://www.columbia lawreview.org/wp-content/uploads/2021/04/Douek-Governing_Online_Speech-from_Posts_As-Trumps_To_Proportionality_And_Probability.pdf. Acesso em: 03/11/2021).

(i) o modelo adequado de responsabilização das plataformas por conteúdo específico publicado por terceiros; (ii) as regras para regulação do poder das plataformas quando estiverem moderando conteúdo com fundamento em termos de uso privados; e (iii) os deveres mínimos de moderação de conteúdo que represente ameaças concretas à democracia e/ou à própria liberdade de expressão.⁴⁶⁰

Para Balkin, o objetivo de uma regulamentação “não é atingir uma neutralidade ilusória na moderação de conteúdo de mídia social. Em vez disso, o objetivo é moldar a organização e os incentivos da indústria para melhor atingir os fins públicos.”⁴⁶¹, traçando, ainda, dois objetivos primordiais, “aumentar o número de participantes para que possa haver muitas empresas, comunidades, recursos e políticas editoriais diferentes”,⁴⁶² e “dar às empresas de mídia social incentivos para se profissionalizar e assumir a responsabilidade pela saúde da esfera pública.”⁴⁶³

Reafirma-se o fato do debate possuir maior profundidade, demandando, desta forma, mais esforços para completá-lo; contudo, podemos expor, minimamente, a fim de conceder ampliação ao tema, os aspectos evidenciados nessa pesquisa.

3.2. Moderação de conteúdo: indo além da remoção

A problemática transfigura-se em: *é possível as redes sociais removerem conteúdos de usuários? Se sim, é o único recurso para corrigir violações as suas regras?* Esse questionamento poderia ser o quinto tópico, o qual se relaciona com os demais tópicos apresentados no item anterior, pois os remédios/recursos de correção de uma violação dependem da legitimidade do sistema de moderação de conteúdo interligado às suas regras e da determinação da temática vista como violação.

⁴⁶⁰ BARROSO, Luna van Brussel. **Liberdade de expressão e democracia na Era Digital**: o impacto das mídias sociais no mundo contemporâneo. Belo Horizonte: Fórum, 2022. p. 229.

⁴⁶¹ Tradução livre do original: “[...] is not to achieve an illusory neutrality in social media content moderation. Rather, the goal is to shape the organization and incentives of the industry to better achieve public ends.” (BALKIN, Jack M., How to Regulate (and Not Regulate) Social Media. **Journal of Free Speech Law**. v. 1, p.71-96, Knight Institute Occasional Paper Series, No. 1, Yale Law School, Public Law Research Paper. 2021. p. 90. Disponível em: <https://ssrn.com/abstract=3484114>. Acesso em: 10/12/2021).

⁴⁶² Tradução livre do original: “[...] increase the number of players so there can be many different companies, communities, affordances, and editorial policies.” (BALKIN, Jack M., How to Regulate (and Not Regulate) Social Media. **Journal of Free Speech Law**. v. 1, p.71-96, Knight Institute Occasional Paper Series, N. 1, Yale Law School, Public Law Research Paper. 2021. Disponível em: <https://ssrn.com/abstract=3484114>. Acesso em: 10/12/2021. p. 90)

⁴⁶³ Tradução livre do original: “[...] give social media companies incentives to professionalize and take responsibility for the health of the public sphere.” (BALKIN, Jack M., How to Regulate (and Not Regulate) Social Media. **Journal of Free Speech Law**. v. 1, p.71-96, Knight Institute Occasional Paper Series, N. 1, Yale Law School, Public Law Research Paper. 2021. Disponível em: <https://ssrn.com/abstract=3484114>. Acesso em: 10/12/2021. p. 90)

Portanto, se o arranjo da moderação de conteúdo não possui força de validação, os remédios resultantes dele não serão legítimos. Contudo, o estudo isolado dos tratamentos, como explica Eric Goldman, favorece a análise dos medicamentos alternativos, a qual, sem essa desassociação, ficaria ofuscada. Ademais, auxilia na obtenção de informações mais detalhadas desse novo mecanismo sem demandar a realização de uma expansão de aprendizado sobre a regulação de conteúdos.⁴⁶⁴

Antes de explorarmos os remédios alternativos, objeto do próximo item (3.3), torna-se importante intensificar a reflexão sobre os padrões de recursos preventivos, alimentando o entendimento sobre medicamentos além da remoção.

Land e Hamilton, a partir de exames sobre a imperatividade da expansão do *kit* de ferramentas para responder ao discurso de ódio nas redes sociais, defendem, de igual modo, a impreterível atuação além da remoção. Afirmando, portanto: “os desenvolvimentos tecnológicos ultrapassaram amplamente as ferramentas jurídicas e políticas existentes para responder ao uso de dispositivos de comunicação de massa para instigar ou perpetuar as violações dos direitos humanos.”⁴⁶⁵ Ademais, devido ao intenso discurso de ódio viralizado nos serviços de *Internet*, as “plataformas de mídia social como *Facebook* e *Twitter* operaram principalmente em modo reacionário – buscando retirar conteúdo ou banir usuários depois que as preocupações foram ao ar em fóruns públicos de alto nível. Isso, afirmamos, é insuficiente.”⁴⁶⁶

Ressalta-se a preocupação em analisar as consequências da violação de uma norma das plataformas por algum tema *online*, ou atividade de usuário. Esse fator é determinado pela doutrina como remédios às consequências “*ex post*”, devido à aplicação desses recursos almejarem punir e prevenir a violação da regra, as quais podem ser chamadas, também, de sanções ou punições; contudo, seria um termo impreciso, pois eliminaria a análise de remédios sem o caráter punitivo.

⁴⁶⁴ GOLDMAN, Eric. Content Moderation Remedies. **Michigan Technonology Law Review**. V. 28, n. 1. p. 1-59, 2021. Disponível em: <https://repository.law.umich.edu/mtlr/vol28/iss1/2>. Acesso em: 02/11/2021. p. 9.

⁴⁶⁵ Tradução livre do original: “Technological developments have largely outpaced existing legal and political tools for responding to the use of mass communications devices to instigate or perpetuate human rights violations.” (LAND, Molly K.; HAMILTON, Rebeca J. Beyond takedown: Expanding the tool kit for responding to online hate. In: DOJČINOVIĆ, Predrag (Ed.). **Propaganda and International Criminal Law: Fron cognition to Criminality**. 1 ed. London: Routledge, 2019. Disponível em: <https://doi.org/10.4324/978042944369>. Acesso em: 01/12/2021).

⁴⁶⁶ Tradução livre do original: “[...] social media platforms like Facebook and Twitter have mostly operated in reactionary mode – seeking to take down content or ban users after concerns have been aired in high-profile public forums. This, we contend, is insufficient.” (LAND, Molly K.; HAMILTON, Rebeca J. Beyond takedown: Expanding the tool kit for responding to online hate. In: DOJČINOVIĆ, Predrag (Ed.). **Propaganda and International Criminal Law: Fron cognition to Criminality**. 1 ed. London: Routledge, 2019. Disponível em: <https://doi.org/10.4324/978042944369>. Acesso em: 01/12/2021).

Kate Klonick explora essa definição ao argumentar os diversos níveis contidos na moderação de conteúdo, ou seja, salientando a possibilidade da ocorrência de restrições *ex ante* ou, após o acontecimento, a *ex post*.⁴⁶⁷ Inobstante, esse conceito também é abarcado como *reativo* ou *proativo*, cita-se:

Esses métodos podem ser reativos, em que os moderadores avaliam passivamente o conteúdo e atualizam o software somente depois que outros chamam a atenção para o conteúdo, ou proativos, em que equipes de moderadores buscam ativamente o conteúdo publicado para remoção. Além disso, essas decisões podem ser automaticamente feitos por software ou feitos manualmente por humanos.⁴⁶⁸

Reforçando este entendimento, Grimmelmann disserta de forma genérica sobre a moderação da *Internet*, afirmando a operação em múltiplas linhas dos sistemas de normatização, como: automático ou manual, transparente ou secreto, *ex ante* ou *ex post* e centralizado ou descentralizado.⁴⁶⁹

Visto isso, iniciaremos um debate mais aprofundado sobre os remédios *ex post*, ou seja, realizados após a postagem considerada violadora. Assim, o estudo centrar-se-á nas decisões tomadas pelos próprios entes privados, não serão analisadas neste trecho, portanto, decisões de estatais ou do sistema jurídico. Diferentemente do Estado, as plataformas – apesar de estarem mais próximas do “Estado *online*”⁴⁷⁰ – não possuem legitimidade do voto para implementarem normas com poder de polícia, saliente-se, assim, que o mecanismo de responsabilidade dos entes privados é diferente daqueles aplicados autoridades públicas.⁴⁷¹

⁴⁶⁷ KLONICK, Kate. The New Governors: The People, Rules and Processes Governing Online Speech. **Harvard Law Review**. v. 131. p. 1598-1670. 2018. Disponível em: https://harvardlawreview.org/wp-content/uploads/2018/04/1598-1670_Online.pdf. Acesso em: 28/05/2022. p. 1635.

⁴⁶⁸ Tradução livre do original: “These methods can be either reactive, in which moderators passively assess content and update software only after others bring the content to their attention, or proactive, in which teams of moderators actively seek out published content for removal. Additionally, these decisions can be automatically made by software or manually made by humans.” (KLONICK, Kate. The New Governors: The People, Rules and Processes Governing Online Speech. **Harvard Law Review**. v. 131. p. 1598-1670. 2018. Disponível em: https://harvardlawreview.org/wp-content/uploads/2018/04/1598-1670_Online.pdf. Acesso em: 28/05/2022. p. 1635).

⁴⁶⁹ GRIMMELMANN, James. The Virtues of Moderation. **The Yale Journal of Law & Technology**. v. 17, n. 1, p. 42-109, 2015. p. 63. Disponível em: <https://core.ac.uk/download/pdf/72836082.pdf>. ou <https://digitalcommons.law.yale.edu/yjolt/vol17/iss1/2>. Acesso em: 10/01/2022.

⁴⁷⁰ Nesse sentido, Sofia Ranchordás afirma que “as plataformas digitais se comportam como um estado ao lado dos estados tradicionais”, fazendo uma associação com as três características tradicionais de um estado: população (usuários), território (plataformas estruturadas) e organização política (autorregulação) e jurídica (governança privada, a exemplo do Conselho de Supervisão do Facebook).” (RANCHORDAS, Sofia, *Digitale platformen: Globale staten naast de staat? Digital Platforms: Global States beyond the State?* **University of Groningen Faculty of Law Research Paper N. 35/2020**, In Besselink, L. (red) *De economische constitutie. Staatsrechtsconferentie 2019*, University of Amsterdam Press, 2020. Disponível em: <https://ssrn.com/abstract=3663065>. Acesso em: 28/10/2021. p. 13 e seguintes).

⁴⁷¹ GOLDMAN, Eric. Content Moderation Remedies. **Michigan Technology Law Review**. v. 28, n. 1. p. 1-59, 2021. Disponível em: <https://repository.law.umich.edu/mtlr/vol28/iss1/2>. Acesso em: 02/11/2021. p. 10.

Ademais, outra característica dos recursos aplicados pelos entes privados, em detrimento do Estado, é o fato de a maioria dos remédios à disposição dos atores estatais não estarem disponíveis às entidades privadas do serviço de *Internet*, como explica Douek, “embora possa haver sérias consequências de ter conteúdo ou contas removidas das mídias sociais, elas geralmente ficam aquém das consequências da sanção estatal.”⁴⁷²

Por outro lado, os atores privados não podem realizar prisões, confiscar bens, realizar busca e apreensão de crianças etc., pois o poder de polícia só pode ser exercido por ente ou entidade pública, não sendo, dessa forma, delegável à entidade privada. Estão restritos a regular o comportamento do usuário dentro de seu território digital – utilizando o termo de Sofia Ranchordàs – ficando bem mais restrito à sua área de atuação. Outro fator distintivo do poder do Estado em relação à estrutura privada está ancorada nos limites de sua aplicação, pois, os atores particulares podem executar seus remédios ou restrições de forma remota, *online*, não havendo limites físicos ou naturais, possibilitando aos serviços de *Internet* criarem e implementarem soluções, mesmo não havendo precedentes históricos *offline*, por exemplo o remédio *toading*.⁴⁷³

Nessa perspectiva, Perel afirma que a aplicação de soluções tecnológicas deve ser afeita ao terreno privado e não às autoridades públicas, pois, “de fato, com remédios digitais, escolher entre várias possibilidades de implementação e aplicá-las é feito em terreno privado, fora do tribunal, não obstante suas importantes implicações para o estado de direito, direitos humanos e inovação.”⁴⁷⁴

Quanto à aplicação dos remédios na moderação de conteúdo, defendemos que haja maior diversificação entre as soluções existentes, entre elas, a remoção codificada, conhecida como “primária ou exclusiva”, a qual é aplicada pelos reguladores utilizando-se das plataformas *online* e das normas e legislações reguladoras, estabelecendo limites, a fim de

⁴⁷² Tradução livre do original: “[...] while there can be serious consequences from having content or accounts removed from social media, these will usually fall short of the consequences of state sanction.” (DOUEK, Evelyn. Governing online speech: from “posts-as-trumps” to proportionality and probability. **Columbia Law Review**. v. 121, n. 3, p. 759-837, 2021. Disponível em: https://www.columbialawreview.org/wp-content/uploads/2021/04/Douek-Governing_Online_Speech-from_Posts_As-Trumps_To_Proportionality_And_Probability.pdf. Acesso em: 03/11/2021. p. 806-807)

⁴⁷³ Segundo Goldman, consiste na prática de converter um avatar de um jogador em um sapo virtual, restringindo sua funcionalidade no jogo. (GOLDMAN, Eric. Content Moderation Remedies. **Michigan Technology Law Review**. v. 28, n. 1. p. 1-59, 2021. Disponível em: <https://repository.law.umich.edu/mtlr/vol28/iss1/2>. Acesso em: 02/11/2021. p. 11)

⁴⁷⁴ Tradução livre do original: “Indeed, with digital remedies, choosing between various implementation possibilities and applying them is done on private grounds, outside the courthouse, notwithstanding its important implications for the rule of law, human rights, and innovation.” (PEREL, Maayan. Digital Remedies. **Berkeley Technology Law Journal**. v. 35, n. 1, p. 1-52, 2020. Disponível em: https://lawcat.berkeley.edu/record/1174610/files/01_Perel_WEB.pdf. Acesso em: 10/07/2022. p. 39).

respeitar os direitos dos usuários, não violando os direitos relativos à liberdade de expressão dos usuários.

A pesquisa de Haimson, Delmonaco, Nie, e Wegner, demonstrou, ao analisar dados sobre os tipos de usuários de mídia social, o seguinte: há determinado conteúdo e contas removidos mais frequentemente que outros; categorizando ainda os tipos de assuntos e de contas retirados e como a remoção pode se dar de forma diferente, dependendo do grupo no qual as restrições ocorram. Outra importante contribuição da pesquisa diz respeito aos grupos de usuários cujas restrições incidem mais frequentemente, são eles: *conservadores políticos*, *pessoas trans* e *pessoas negras*. Contudo, o conteúdo removido se difere em cada grupo, sendo assim: conservadores “incluía conteúdo ofensivo ou supostamente ofensivo, desinformação, relacionado à Covid, adulto ou discurso de ódio”.⁴⁷⁵ O conteúdo dos usuários transgêneros era frequentemente removido “como adulto, apesar de seguir as diretrizes do site, criticar um grupo dominante ou especificamente relacionado a questões de transgêneros ou questões estranhas”.⁴⁷⁶ Enquanto o conteúdo removido de participantes negros era frequentemente “relacionado à justiça racial ou racismo”.⁴⁷⁷

Perceberam, nesse panorama, que as “remoções de participantes conservadores geralmente envolviam conteúdo nocivo removido de acordo com as diretrizes do site”,⁴⁷⁸ buscando deixar plataforma com informações verídicas e precisas. E o que mais chamou a atenção foram as “remoções de participantes transgêneros e negros, que geralmente envolviam conteúdo relacionado à expressão de suas identidades marginalizadas,”⁴⁷⁹ embora

⁴⁷⁵ HAIMSON, Oliver L.; DELMONACO, Daniel; NIE, Peipei; WEGNER, Andrea. Disproportionate Removals and Differing Content Moderation Experiences for Conservative, Transgender, and Black Social Media Users: Marginalization and Moderation Gray Areas. **Proceeding of the ACM on Human-Computer Interaction**. v. 5, n. CSCW2, Artigo 466, p. 1-35. 2021. Disponível em: <https://doi.org/10.1145/3479610>. Acesso em: 10/07/2022. p. 466:1 e seguintes.

⁴⁷⁶ HAIMSON, Oliver L.; DELMONACO, Daniel; NIE, Peipei; WEGNER, Andrea. Disproportionate Removals and Differing Content Moderation Experiences for Conservative, Transgender, and Black Social Media Users: Marginalization and Moderation Gray Areas. **Proceeding of the ACM on Human-Computer Interaction**. v. 5, n. CSCW2, Artigo 466, p. 1-35. 2021. Disponível em: <https://doi.org/10.1145/3479610>. Acesso em: 10/07/2022. p. 466:1 e seguintes.

⁴⁷⁷ HAIMSON, Oliver L.; DELMONACO, Daniel; NIE, Peipei; WEGNER, Andrea. Disproportionate Removals and Differing Content Moderation Experiences for Conservative, Transgender, and Black Social Media Users: Marginalization and Moderation Gray Areas. **Proceeding of the ACM on Human-Computer Interaction**. v. 5, n. CSCW2, Artigo 466, p. 1-35. 2021. Disponível em: <https://doi.org/10.1145/3479610>. Acesso em: 10/07/2022. p. 466:1 e seguintes.

⁴⁷⁸ HAIMSON, Oliver L.; DELMONACO, Daniel; NIE, Peipei; WEGNER, Andrea. Disproportionate Removals and Differing Content Moderation Experiences for Conservative, Transgender, and Black Social Media Users: Marginalization and Moderation Gray Areas. **Proceeding of the ACM on Human-Computer Interaction**. v. 5, n. CSCW2, Artigo 466, p. 1-35. 2021. Disponível em: <https://doi.org/10.1145/3479610>. Acesso em: 10/07/2022. p. 466:1 e seguintes.

⁴⁷⁹ HAIMSON, Oliver L.; DELMONACO, Daniel; NIE, Peipei; WEGNER, Andrea. Disproportionate Removals and Differing Content Moderation Experiences for Conservative, Transgender, and Black Social Media Users: Marginalization and Moderation Gray Areas. **Proceeding of the ACM on Human-Computer Interaction**. v.

seguissem as regras do serviço de *internet* ou caíram numa zona opaca da moderação de conteúdo e foram removidas.

Verifica-se com isso a moderação de conteúdo concentrada na remoção como um “remédio padrão”, numa ideia binária, além de tornar a moderação menos equitativa e mais desigual para usuários já marginalizados, dessa vez também pela mídia social, onde continuam sendo silenciados, ferindo-lhes a liberdade de expressão.

Como exemplos clássicos desta postura de moderação de conteúdo cuja remoção é privilegiada, a doutrina aponta a previsão numa lei Americana chamada de *The Digital Millenium Copyright Act* – DMCA. Essa lei trata sobre os direitos autorais, e numa atualização da lei para a era digital, instituiu uma solução binária de remoção de conteúdo, propondo o encerramento de conta em qualquer dos dois casos. A primeira delas é conhecida como cláusula “*notice-and-takedown*” ou “notificação e remoção”, nesse caso os proprietários dos direitos autorais notificam os serviços de *uploads* de usuários, supostamente infratores, e os serviços de *Internet* “remove ou desabilita o acesso aos arquivos do usuário, como resposta ao aviso do proprietário”.⁴⁸⁰ A segunda forma é o encerramento da conta, consistindo em identificar reincidentes, por meio da emissão de “*strikes*”, como se fosse um alerta ou uma suspensão para o usuários e, a partir de uma quantidade desses alertas, a conta é encerrada.⁴⁸¹

Logo após a utilização dessa lei pelos norteamericanos, a União Europeia seguiu na mesma linha legislativa e elaborou a “Diretiva de comércio eletrônico”. Por exemplo, em 2017, a Alemanha aprovou o *Netzwerkdurchsetzungsgesetz* – o *Network Enforcement Act*

5, n. CSCW2, Artigo 466, p. 1-35. 2021. Disponível em: <https://doi.org/10.1145/3479610>. Acesso em: 10/07/2022. p. 466:1 e seguintes

⁴⁸⁰ Conforme previsão *Digital Millennium Copyright Act*, § 512, (C) (1) (A) (iii)): “upon obtaining such knowledge or awareness, acts expeditiously to remove, or disable access to, the material;” (Tradução livre do original: “ao obter tal conhecimento ou cconhecimento, age rapidamente para remover ou desabilitar o acesso ao material;” e § 512 (c) (1) (C): “upon notification of claimed infringement as described in paragraph (3), responds expeditiously to remove, or disable access to, the material that is claimed to be infringing or to be the subject of infringing activity.” Tradução livre do original: “após a notificação da violação alegada, conforme descrito no parágrafo (3), responde rapidamente para remover ou desabilitar o acesso ao material que alegadamente infringiu ou foi objeto de atividade infratora.” (EUA, **Digital Millennium Copyright Act**. Pub. L. 105-304, 112 Stat. 2860, 1998. Disponível em: <https://www.congress.gov/105/plaws/publ304/PLAW-105publ304.pdf>. Acesso em: 10/07/2022).

⁴⁸¹ “Os strikes são a maneira usada pela plataforma para penalizar os criadores de conteúdo e fazer valer suas diretrizes de comunidade e publicação. A postagem de materiais protegidos por direitos autorais, por exemplo, pode gerar penalização do tipo caso haja uma solicitação de remoção por parte do detentor do copyright. Vídeos com conteúdo sexual e explícito, imagens violentas ou que contenham abuso ou discurso de ódio também vão contra os termos de uso da plataforma”. (DEMARTINI, Felipe. Youtube muda sistema de strikes e avisos por quebra de diretrizes. Canaltech. 2019. Disponível em: <https://canaltech.com.br/internet/youtube-muda-sistema-de-strikes-e-avisos-por-quebra-de-diretrizes-133113/>. Acesso em: 11/07/2022).

(“NetzDG”), exigindo dos serviços de *Internet* que “removam ou bloqueiem o acesso” de lista de categorias de conteúdo considerados ilegais, num prazo bastante exíguo.⁴⁸²

Os exemplos já apresentados são oriundos de legislações internacionais oriundos de entes públicos. Os próximos dois foram extraídos de Organizações da Sociedade Civil, o primeiro: *Os princípios de Manila, sobre responsabilidade de intermediário*,⁴⁸³ possuindo como objetivo “proteger a liberdade de expressão e criar um ambiente propício para a inovação, que equilibre as necessidades dos governos e outras partes interessadas, grupos da sociedade civil de todo o mundo [...]”.

Os princípios buscam evitar a remoção injustificadas ou injustas de conteúdo postados por usuários, apesar de utilizar com frequência o termo “restrições de conteúdo”, em vez de “remoções de conteúdo”. No item IV dos Princípios, estipula-se o seguinte: “Leis e ordens e práticas de restrição de conteúdo devem cumprir os testes de necessidade e proporcionalidade”, buscando-se antever as seguintes situações:

- a) Qualquer restrição de conteúdo deve ser limitada ao conteúdo específico em questão
- b) Na restrição de conteúdo, devem ser adotados os meios técnicos menos restritivos.
- c) Se o conteúdo for restrito porque é ilegal em uma determinada região geográfica e se o intermediário oferecer um serviço geograficamente diversificado, o escopo geográfico da restrição de conteúdo deve ser limitado.
- d) Se o conteúdo for restringido devido à sua ilegalidade por um período limitado, a restrição não deve durar além dessa duração, e a ordem de restrição deve ser revisada periodicamente para garantir que permaneça válida.⁴⁸⁴

O segundo exemplo, também extraído de Organizações da Sociedade Civil, são os *Princípios de Santa Clara*,⁴⁸⁵ originando-se de uma publicação acadêmica sobre transparência

⁴⁸² Conferir: EIFERT, Martin. A Lei Alemã para a Melhoria da Aplicação da Lei nas Redes Sociais (NetzDG) e a Regulação da Plataforma. In: ABOUD, Georges; NERY JÚNIOR, Nelson; CAMPOS, Ricardo (Coord.). **Fake news e regulação**. 2 ed. São Paulo: Thomson Reuters Brasil, 2020.

⁴⁸³ Tradução livre do original: “[...] he aim of protecting freedom of expression and creating an enabling environment for innovation, which balances the needs of governments and other stakeholders, civil Society groups from around the world.” (**Princípios de Manila sobre Responsabilidade Intermediária**, Versão 1.0, 2015. p. 1. Disponível em: https://www.eff.org/files/2015/10/31/manila_principles_1.0.pdf. Acesso em: 12/07/2022).

⁴⁸⁴ Tradução livre do original: “IV. Laws and content restriction orders and practices must comply with the tests of necessity and proportionality.

Laws, orders and practices restricting content must be necessary and proportionate in a democratic society:

- a) Any restriction of content should be limited to the specific content at issue.
- b) When restricting content, the least restrictive technical means must be adopted.
- c) If content is restricted because it is unlawful in a particular geographical region, and if the intermediary offers a geographically variegated service, then the geographical scope of the content restriction must be so limited.
- d) If content is restricted owing to its unlawfulness for a limited duration, the restriction must not last beyond this duration, and the restriction order must be reviewed periodically to ensure it remains valid.” (**Princípios de Manila sobre Responsabilidade Intermediária**, Versão 1.0, 2015. Disponível em: https://www.eff.org/files/2015/10/31/manila_principles_1.0.pdf. Acesso em: 12/07/2022. p. 4)

e responsabilidade na moderação de conteúdo. Desde 2018, grandes empresas endossaram os Princípios – dentre elas *Apple*, *Facebook* (*Meta*), *Google*, *Reddit*, *Twitter* e *Github*. Atualmente, houve uma nova coalizão entre as organizações, os defensores e os especialistas acadêmicos, e durante os anos 2020 e 2021, desenvolveram, ou melhor, atualizaram uma segunda iteração dos Princípios de Santa Clara.

Nessa segunda versão, os Princípios Fundamentais e Operacionais são divididos. Sendo os Princípios Fundamentais constituintes de ideários mais abrangentes e transversais, devendo ser considerados em todas as empresas, nas quais haja qualquer tipo de moderação de conteúdo. De outro modo, os Princípios Operacionais estabelecem arranjos para empresas maiores ou mais experientes quanto ao processo de moderação de conteúdo.⁴⁸⁶

Nessa segunda iteração, a transparência é exigida em relação aos “conteúdo” e “ação” realizados por uma empresa. O termo “conteúdo” abrange todo teor gerado pelo usuário – pago ou não - incluindo publicidade. Os termos “ação” e “acionado” referem-se a qualquer ato de execução realizada por uma empresa, em relação ao conteúdo ou à conta de um usuário devido à não conformidade com suas regras ou políticas, dentre elas: a remoção do teor da postagem, *downranking* algorítmico de conteúdo e a suspensão – temporária ou permanente – de contas. Buscando, com isso, apoiar as empresas a cumprirem as suas responsabilidades de respeitar os direitos humanos bem como aumentarem suas prestações de contas e para auxiliar os defensores dos direitos humanos em suas atividades.⁴⁸⁷

Demonstrou-se, nesse tópico, que os reguladores e estudiosos tratam o remédio de remoção como uma solução principal na moderação de conteúdo. Isso é motivado, muitas vezes, pela facilidade de sua aplicação, como os reguladores podem descrever sem maiores esforços esse remédio. Ademais, com a remoção evita a continuidade de violações legais, bem como facilita na verificação das medidas adotadas. Contudo, como explica Gillespie, a remoção é uma “abordagem mais dura, em termos de suas consequências”,⁴⁸⁸ gerando efeitos

⁴⁸⁵ “Esses princípios, nomeados em homenagem ao local da reunião inicial do grupo em Santa Clara, CA, representam recomendações para os passos iniciais que as empresas envolvidas na moderação de conteúdo devem tomar para fornecer um devido processo significativo aos palestrantes afetados e garantir melhor que a aplicação de suas diretrizes de conteúdo seja justa, imparcial, proporcional e respeitosa dos direitos dos usuários. Esta foi a primeira iteração dos Princípios de Santa Clara.” (**The Santa Clara Principles: On Transparency and Accountability in Content Moderation**. Disponível em: <https://santaclaraprinciples.org/>. Acesso em: 12/07/2022).

⁴⁸⁶ **The Santa Clara Principles: On Transparency and Accountability in Content Moderation**. Disponível em: <https://santaclaraprinciples.org/>. Acesso em: 12/07/2022.

⁴⁸⁷ **The Santa Clara Principles: On Transparency and Accountability in Content Moderation**. Disponível em: <https://santaclaraprinciples.org/>. Acesso em: 12/07/2022.

⁴⁸⁸ Tradução livre do original: “It is the harshest approach, in terms of its consequences.” (GILLESPIE, Tarleton. **Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media**. Yale University Press/New Haven & London. 2018. Disponível em: https://www.researchgate.net/publication/327186182_Custodians_of_the_internet_Platforms_content_moderatio

colaterais, pois torna invisível o conteúdo ou o usuário. E arremata, “a remoção é um instrumento contundente, uma determinação de tudo ou nada, removendo esse conteúdo para todos, não apenas para aqueles que foram ofendidos.”⁴⁸⁹

Além disso, pode acarretar outros problemas, conforme expõe Goldman: a) as remoções eliminam totalmente os registros da violação, com isso, podem excluir elementos de importância histórica ou, em casos de conduta criminosa, perder as provas; b) quando há a exclusão de itens de um conteúdo, deixam o restante fora de contexto ou com comentários vazios desconexos; c) as remoções também excluem os *links* de entrada, impossibilitando que outro usuário siga os *links*; d) quando há remoção de contas, é inevitável o dano colateral, como: - exclusão de uma conta, necessariamente remove todo conteúdo existente nela, inclusive, outros conteúdos que não violam as regras; - a exclusão de um conta, excluirá também o login em outras plataformas ou serviços de *internet*, vinculados aquela conta, ou outros serviços da mesma plataforma.⁴⁹⁰

Com efeito, reconhecemos que as remoções integram o *kit* de ferramentas de correção dos serviços de *Internet*. Contudo, espera-se não ser a única ferramenta ou a mais importante do leque existente de remédios alternativos à remoção, pois, é possível balancear os benefícios e os danos na permanência de determinada postagem. Destarte, há uma promoção da liberdade de expressão e ainda corrige o conteúdo violador das regras da plataforma.

Noutra visão, os remédios de não-remoção podem auxiliar na modificação da concepção comum, de que os serviços de *internet* praticam “censura” com seus usuários, ou os conteúdos são removidos ou suas contas são excluídas. Contribuindo para alterar a visão de “controle total” das plataformas, pois, há meios para buscarem a correção e a permanência do conteúdo.

Depreende-se, com isso, uma importante constatação – sendo este o cerne do nosso trabalho: “criar melhores equilíbrios entre liberdade de expressão e policiamento de conteúdo,

n_and_the_hidden_decisions_that_shape_social_media. Acesso em: 23/11/2021. p. 176). (GILLESPIE, Tarleton. **Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media.** Yale University Press/New Haven & London. 2018. Disponível em: https://www.researchgate.net/publication/327186182_Custodians_of_the_internet_Platforms_content_moderation_and_the_hidden_decisions_that_shape_social_media. Acesso em: 23/11/2021. p. 176).

⁴⁸⁹ Tradução livre do original: “Removal is a blunt instrument, an all- or- nothing determination, removing that content for everyone, not just for those who were offended.” (GILLESPIE, Tarleton. **Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media.** Yale University Press/New Haven & London, 2018. Disponível em: https://www.researchgate.net/publication/327186182_Custodians_of_the_internet_Platforms_content_moderation_and_the_hidden_decisions_that_shape_social_media. Acesso em: 23/11/2021. p.176).

⁴⁹⁰ GOLDMAN, Eric. Content Moderation Remedies. **Michigan Technology Law Review**. v. 28, n. 1. p. 1-59, 2021. Disponível em: <https://repository.law.umich.edu/mtlr/vol28/iss1/2>. Acesso em: 02/11/2021. p. 21.

a implantação mais ampla de remédios de não remoção pode reduzir a ira pública e a temperatura dos debates sobre políticas.”⁴⁹¹

Ademais, a visão não binária, com resposta criativa cuja remoção não seja a única opção, atende às comunidades *online* com públicos diversos, com características distintas e peculiares os quais não se satisfazem apenas com retorno aos seus questionamentos de forma binária.

Desta forma, se as remoções fossem um recurso exclusivo ou principal em todos os serviços de *internet*, certamente essa diversidade iria ser restringida por um remédio padrão, que não serve para manter a natureza distinta dos usuários.

Por outro lado, um *kit* de ferramentas de remédios variados permitiria: “os serviços da *Internet* refinem e otimizem suas abordagens de moderação de conteúdo para melhor atender às necessidades específicas de sua comunidade.”⁴⁹² Outro ponto positivo, na diversificação de remédios, pode representar uma “estratégia” para se tornar um ponto diferencial aumentando a competitividade em detrimento às demais plataformas. Ou seja, as plataformas concorrentes podem adotar estratégias diferentes e permitir que os públicos decidam qual abordagem, tipo de comunidade ou recursos desejam. Com isso, contribui para aumentar a concorrência no mercado e ajudar os serviços a atender melhor seus públicos.⁴⁹³

3.3. Remédios alternativos de moderação

É possível compreender o motivo de a maioria das plataformas *online* possuírem o comportamento de não divulgar, precisamente, como constroem e chegam às suas decisões de moderação de conteúdo. Com isso, adotam a remoção de conteúdo como principal ferramenta, sem fornecer ao usuário uma explicação ou justificativa.⁴⁹⁴ Assim, essa deficitária transparência no processo decisório das plataformas *online* induz a uma ausência de dados de

⁴⁹¹ Tradução livre do original: “By creating better balances between free expression and content policing, wider deployment of non-removal remedies may reduce public ire and the temperature of the policy debates.” (GOLDMAN, Eric. Content Moderation Remedies. **Michigan Technology Law Review**. v. 28, n. 1. p. 1-59, 2021. Disponível em: <https://repository.law.umich.edu/mtlr/vol28/iss1/2>. Acesso em: 02/11/2021. p. 23).

⁴⁹² Tradução livre do original: “[...] Internet services refine and optimize their content moderation approaches to best cater to their specific community’s needs.” (GOLDMAN, Eric. Content Moderation Remedies. **Michigan Technology Law Review**. v. 28, n. 1. p. 1-59, 2021. Disponível em: <https://repository.law.umich.edu/mtlr/vol28/iss1/2>. Acesso em: 02/11/2021. p. 23)

⁴⁹³ GOLDMAN, Eric. Content Moderation Remedies. **Michigan Technology Law Review**. v. 28, n. 1. p. 1-59, 2021. Disponível em: <https://repository.law.umich.edu/mtlr/vol28/iss1/2>. Acesso em: 02/11/2021. p. 23.

⁴⁹⁴ WAGNER, Ben; KÜBLER, Johanne; PÍRKOVÁ, Eliška; GSENGER, Rita; FERRO, Carolina. **Reimagining Content Moderation and Safeguarding Fundamental Rights: A Study on Community-Led Platforms**. The Green/EFA in the European Parliament. 2021. p. 24. Disponível em: https://www.greens-efa.eu/files/assets/docs/alternative_content_web.pdf. Acesso em: 15/07/2022.

boa qualidade⁴⁹⁵, dificultando o desenvolvimento de pesquisas sobre o tema. Jhaver, Appling, Filbert e Bruckman reforçam como essa opacidade nos arranjos de moderação de conteúdo tem fortalecido a construção de “teorias populares”⁴⁹⁶ pelos usuários, a respeito de como e por qual motivo o seu conteúdo foi removido.

Nessa esteira, as pesquisas têm demandado esforços para estudar estratégias, apresentando soluções alternativas, a fim de as plataformas não ficarem engessadas, fazendo mais que apenas remover o conteúdo como “sanção”. Dessa forma, o teor de postagens indesejáveis ou maus usuários poderão ser punidos como meio de garantir um espaço digital saudável, aberto e democrático. Dentre elas, apontamos a pesquisa de Jhaver, Appling, Filbert e Bruckman a qual tenta aumentar a transparência no processo de moderação. O estudo utilizou a plataforma *Reddit*, influenciando os moderadores a fornecerem explicações sobre as remoções aos usuários, em uma postagem curta – chamada “*flairing*”⁴⁹⁷.

Nesse raciocínio, nosso estudo apresenta um *kit* de ferramentas de remédios expandidos,⁴⁹⁸ como soluções alternativas para violações de regras *online*, ou, como prefere Goldman e Miers: “regras da casa”,⁴⁹⁹ transcendendo a ideia de simplesmente remover o

⁴⁹⁵ SUZOR, Nicolas; VAN GEELEN, Tess; WEST, Sarah Myers. Evaluating the Legitimacy of Platform Governance: A Review of Research and a Shared Research Agenda. **International Communication Gazette**. v. 80, n. 4, p. 385–400. 2018. p. 395. Disponível em: doi: 10.1177/1748048518757142 ou <https://osf.io/preprints/lawarxiv/jak8d/>. Acesso em: 10/06/2022.

⁴⁹⁶ JHAVER, Shagun; APPLING, Darren Scott; FILBERT, Eric; BRUCKMAN, Amy. “Did You Suspect the Post Would be Removed?”: Understanding User Reactions to Content Removals on Reddit. **Proceedings of the ACM on Human-Computer Interaction**. v. 3, CSCW, Article 192. ACM, New York, NY. 2019. p. 192:3. Disponível em: <https://doi.org/10.1145/3359294> ou <https://dl.acm.org/doi/pdf/10.1145/3359294>. Acesso em: 15/07/2022.

⁴⁹⁷ Importante consignar, umas das conclusões desse estudo, conferir mais atenção à educação, em vez da remoção: “[...] uma ênfase na educação, em vez da remoção, pode melhorar a visão dos usuários em relação à comunidade e incentivá-los a participar de forma construtiva. Projetar ferramentas automatizada que possam fornecer explicações de remoção em cenários específicos pode ajudar os recém-chegados a se familiarizarem com as normas da comunidade sem aumentar indevidamente o trabalho dos moderadores. A criação de diretrizes da comunidade que sejam claras e fáceis de seguir pode melhorar a percepção de justiça dos usuários com moderação.” Tradução livre do original: “an emphasis on education, rather than removal, may improve users’ outlook towards the community and encourage them to participate constructively. Designing automated tools that can provide removal explanations in specific scenarios can help newcomers become familiar with the norms of community without unduly increasing the work of moderators. Creating Community guidelines that are clear and easy to follow can improve users’ perceptions of fairness in moderation.” (JHAVER, Shagun; APPLING, Darren Scott; FILBERT, Eric; BRUCKMAN, Amy. “Did You Suspect the Post Would be Removed?”: Understanding User Reactions to Content Removals on Reddit. **Proceedings of the ACM on Human-Computer Interaction**. v. 3, CSCW, Article 192. ACM, New York, NY. 2019. p. 192:27. Disponível em: <https://doi.org/10.1145/3359294> ou <https://dl.acm.org/doi/pdf/10.1145/3359294>. Acesso em: 15/07/2022).

⁴⁹⁸ Conferir: LAND, Molly K.; HAMILTON, Rebeca J. Beyond takedown: Expanding the tool kit for responding to online hate. In: DOJČINOVIĆ, Predrag (Ed.). **Propaganda and International Criminal Law: From cognition to Criminality**. 1 ed. London: Routledge, 2019. Disponível em: <https://doi.org/10.4324/978042944369>. Acesso em: 01/12/2021; e GOLDMAN, Eric. Content Moderation Remedies. **Michigan Technology Law Review**. v. 28, n. 1. p. 1-59, 2021. Disponível em: <https://repository.law.umich.edu/mtlr/vol28/iss1/2>. Acesso em: 02/11/2021.

⁴⁹⁹ GOLDMAN, Eric; MIERS, Jess. Online Account Terminations/Content Removals and the Benefits of Internet Services Enforcing Their House Rules. **Journal of Free Speech Law**. v. 1, n. 191, Santa Clara

conteúdo, motivo no qual, nas palavras de Wagner, Kübler, Pírková, Gsenger e Ferro, “é simplesmente um ‘Band-Aid’ para um problema já existente”.⁵⁰⁰ Considerando tratar-se de tecnologia, os remédios podem, de várias (ou infinitas) formas.

Abordaremos a taxonomia de opções de remédios proposta por Goldman⁵⁰¹ na qual, basicamente, dividindo-se em cinco categorias: 1. ações em face de parte de conteúdo individuais; 2. ações contra conta *online*; 3. ações que visa reduzir a visibilidade das violações, tanto para conteúdo individuais ou conta inteira; 4. ações impondo consequências financeiras por violações, tanto em face de conteúdo individuais ou conta inteira; e 5. outras ações que não se enquadram nas categorias anteriores.⁵⁰²

Para melhor ilustrar, colacionamos o gráfico elaborado por Goldman reunindo-se todas as ações.⁵⁰³

University Legal Research Paper, 2021. Disponível em: <https://ssrn.com/abstract=3911509>. Acesso em: 03/11/2021.

⁵⁰⁰ Tradução livre do original: “[...] it is simply a ‘Band-Aid’ for an already existing problem.” (WAGNER, Ben; KÜBLER, Johanne; PÍRKOVÁ, Eliška; GSENGER, Rita; FERRO, Carolina. **Reimagining Content Moderation and Safeguarding Fundamental Rights: A Study on Community-Led Platforms**. The Green/EFA in the European Parliament. 2021. p. 29. Disponível em: https://www.greens-efa.eu/files/assets/docs/alternative_content_web.pdf. Acesso em: 15/07/2022).

⁵⁰¹ GOLDMAN, Eric. Content Moderation Remedies. **Michigan Technology Law Review**. v. 28, n. 1. p. 1-59, 2021. Disponível em: <https://repository.law.umich.edu/mlr/vol28/iss1/2>. Acesso em: 02/11/2021. p. 23-24.

⁵⁰² Além dessa classificação, a doutrina aponta outras, que incluem ações antes da violação. A respeito conferir: KRAUT, Robert E.; RESNICK, Paul (Ed.). **Building successful online communities: Evidence-based social design**. Cambridge, Mit Press, 2012. Disponível em: <https://www.jstor.org/stable/j.ctt5hhgvw>. Acesso em: 10/07/2022.

⁵⁰³ GOLDMAN, Eric. Content Moderation Remedies. **Michigan Technology Law Review**. v. 28, n. 1. p. 1-59, 2021. Disponível em: <https://repository.law.umich.edu/mlr/vol28/iss1/2>. Acesso em: 02/11/2021. p. 24.

REGULAMENTO DE CONTEÚDO	REGULAMENTO DE CONTA	REGULAMENTO DE VISIBILIDADE (por ato ou item)	MONETÁRIO (por ato ou item)	OUTROS ⁵⁰⁴
- remover conteúdo - suspender conteúdo - recolocar o conteúdo - editar/redigir conteúdo - aviso intersticial - adicionar legenda de aviso - adicionar contra-fala - desabilitar comentários	- terminar conta - suspender conta - suspender direitos de postagem - remover distintivos de credibilidade - reduzir níveis de serviço (dados, velocidade, etc) - Envergonhar (<i>Shadowban</i>)	- Shadowban - remover do índice de pesquisa externo - Não siga <i>link</i> de autores - Remover do índice de pesquisa interna - rebaixar a visibilidade da pesquisa interna - sem auto-sugestão - nenhuma/reduzida promoção interna - nenhum/reduzido links de navegação - viralidade reduzida - age-gate - exibir apenas para leitores conectados	- lucros acumulados perdidos - encerrar ganhos futuros (por item ou conta) - suspender ganhos futuros (por item ou conta) - multar autor/impôr indenização	- educar os usuários - atribuir strikes/warnings - saída/desmascar (outing/unmasking) - areportar à aplicação da lei - colocar usuário/conteúdo na lista de bloqueio - serviço comunitário - "justiça restaurativa"/des culpas

Para melhor elucidar cada um dos recursos, passaremos a examinar os principais remédios de cada categoria. O primeiro, refere-se a:

1. *Regulação de conteúdo*: possui oito recursos quanto a conteúdos individuais:

1.1. *Remover o conteúdo*: esse recurso é o tem-se tornado padrão pelos moderadores de conteúdo das plataformas e, uma vez executado, não há como restaurar. A execução de uma remoção pode englobar toda uma rede ou apenas determinadas áreas geográficas determinadas pelo moderador o qual também pode optar por remover partes da rede.

Quanto à remoção permanente de conteúdo, Kattemann e Tiedeke indagam: os usuários podem processar plataformas para restabelecer conteúdo removido? Realizam uma

⁵⁰⁴ Tradução livre do original:

Content Regulation	Account Regulation	Visibility Reductions (by acct or item)	Monetary (by acct or item)	Other
• Remove content • Suspend content • Relocate content • Edit/redact content • Interstitial warning • Add warning legend • Add counterspeech • Disable comments	• Terminate account • Suspend account • Suspend posting rights • Remove credibility badges • Reduced service levels (data, speed, etc.) • Shaming	• Shadowban • Remove from external search index • Nofollow authors' links • Remove from internal search index • Downgrade internal search visibility • No auto-suggest • No/reduced internal promotion • No/reduced navigation links • Reduced virality • Age-gate • Display only to logged-in readers	• Forfeit accrued earnings • Terminate future earning (by item or account) • Suspend future earning (by item or account) • Fine author/impose liquidated damages	• Educate users • Assign strikes/warnings • Outing/unmasking • Report to law enforcement • Put user/content on blocklist • Community service • "Restorative justice"/apology

(GOLDMAN, Eric. Content Moderation Remedies. *Michigan Technology Law Review*. v. 28, n. 1. p. 1-59, 2021. p. 24. Disponível em: <https://repository.law.umich.edu/mtlr/vol28/iss1/2>. Acesso em: 02/11/2021).

análise do ponto de vista do “*must carry*”⁵⁰⁵, sob as lentes de jurisprudência dos EUA e da Alemanha, apresentando a conclusão de que nos EUA e na Alemanha, é possível a restrição de conteúdo em sua plataforma através dos termos de serviço. Contudo, dependendo da importância de uma comunicação feita e do poder de mercado, “os serviços de redes sociais na Alemanha enfrentam restrições ao limitar o acesso à plataforma, suspendendo usuários ou cancelando contratos de acesso ao perfil via o conceito de efeito indireto de terceiros dos direitos fundamentais.”⁵⁰⁶ Enquanto a lei e a jurisprudência dos Estados Unidos “em geral relutam em reconhecer deveres baseados em direitos fundamentais para intermediários privados. Como *Knight v. Trump* [...]”⁵⁰⁷

Nessa mesma linha de pesquisa, Edoardo Celeste identificou que esse fenômeno de exclusão da mídia social, configurado quando o usuário é impedido de usar as plataformas de

⁵⁰⁵ Kattemann e Tiedeke, esclarecem: “‘Must carry’ do ponto de vista dos EUA originou-se de um conjunto de regras instituídas pela Federal Communications Commission (FCC) em 1965. Originalmente, ‘must carry’ era uma reivindicação, que foi apresentada com o objetivo de ‘preservar’ uma multiplicidade de emissoras» (Turner II v. FCC, 1997) e que obrigava as redes de televisão por cabo a transmitir determinados programas (locais). No contexto do direito das comunicações, ‘must carry’ não é estranho ao ordenamento jurídico europeu e, em particular, alemão. Faz parte das obrigações estatutárias de transmissão que se aplicam sob o Tratado de Radiodifusão Interestadual Alemão (RStV) aos chamados provedores de plataforma, em particular aos operadores de rede a cabo. Mas o contexto em que os argumentos ‘devem levar’ são apresentados tem expandido nos últimos anos. Palestrantes em plataformas online nos EUA tentaram usar a ideia por trás do ‘must carry’ em um contexto diferente. Eles tentaram forçar plataformas de propriedade privada (mídia social) a veicular (ou seja, publicar) seu discurso. Em sua opinião, ‘uma entidade privada torna-se um ator estatal através de sua operação’ da propriedade privada como ‘um fórum público para o discurso’ (Cyber Promotions v. America Online, 1996). A abordagem das obrigações das plataformas digitais privadas em relação ao discurso a ser veiculado difere significativamente entre os EUA e a Alemanha”. Tradução livre do original: ‘Must carry’ from a US perspective originated from a set of rules instituted by the Federal Communications Commission (FCC) in 1965. Originally, ‘must carry’ was a claim, which was brought forward with the aim ‘to preserve[ing] a multiplicity of broadcasters’ (Turner II v. FCC, 1997) and that obliged cable television networks to carry particular (local) programmes. In the context of communications law, ‘must carry’ is not alien to the European and in particular Germany legal order. It is part of the statutory broadcasting obligations which apply under the German Interstate Broadcasting Treaty (RStV) to so-called platform providers, in particular cable network operators. But the context in which ‘must carry’ arguments are put forward has expanded in the last years. Speakers on online platforms in the US have tried to make use of the idea behind ‘must carry’ in a different context. They have tried to force privately owned (social media) platforms to carry (i.e. publish) their speech. In their view ‘a private entity becomes a state actor through its operation’ of the private property as ‘a public forum for speech’ (Cyber Promotions v. America Online, 1996). The approach to obligations of private digital platforms regarding speech to be carried differs significantly between the US and Germany.” (KETTEMANN, Matthias C.; TIEDEKE, Anna Sophia. Back up: Can users sue platforms to reinstate deleted content? **Internet Policy Review**. Alexander von Humboldt Institute for Internet and Society, Berlim, v. 9, n. 2, p. 1-20. 2020. p. 18-19. Disponível em: <http://dx.doi.org/10.14763/2020.2.1484>. Acesso em: 18/07/2022).

⁵⁰⁶ Tradução livre do original: “[...] social network services in Germany face restrictions in limiting access to the platform by suspending users or cancelling profile access contracts via the concept of indirect third-party effect of fundamental rights.” (KETTEMANN, Matthias C.; TIEDEKE, Anna Sophia. Back up: Can users sue platforms to reinstate deleted content? **Internet Policy Review**. Alexander von Humboldt Institute for Internet and Society, Berlim, v. 9, n. 2, p. 1-20. 2020. p. 11. Disponível em: <http://dx.doi.org/10.14763/2020.2.1484>. Acesso em: 18/07/2022).

⁵⁰⁷ Tradução livre do original: “US law and jurisprudence in general is reluctant to recognise fundamental right-based duties for private intermediaries. As *Knight v. Trump* [...]” (KETTEMANN, Matthias C.; TIEDEKE, Anna Sophia. Back up: Can users sue platforms to reinstate deleted content? **Internet Policy Review**. Alexander von Humboldt Institute for Internet and Society, Berlim, v. 9, n. 2, p. 1-20. 2020. p. 11. Disponível em: <http://dx.doi.org/10.14763/2020.2.1484>. Acesso em: 18/07/2022).

mídias sociais específicas ou parte delas, provoca profundas implicações nos direitos fundamentais.⁵⁰⁸ Assim, reforça a ideia de não padronizar a remoção como meio mais eficaz na moderação. Isso por que as restrições de acesso de usuário as plataformas *online*, ou uma parcela delas, impede o indivíduo de exercer seus direitos fundamentais, sobretudo na era digital, que atingiu um nível tão elevado, sendo esta a regra nos últimos anos.⁵⁰⁹

Cumprе consignar a importante observação realizada por Giovanni De Gregorio, sobre a necessidade das plataformas *online*, quando decidirem por restringir conteúdos, assegurarem a possibilidade do seu restabelecimento.⁵¹⁰ Desta forma, caso o usuário discorde da primeira decisão e recorra ao mecanismo de reparação estabelecido pela plataforma, “é necessário que o conteúdo anteriormente removido continue disponível caso as plataformas *online* revejam a sua primeira decisão. Portanto, a reintegração de conteúdo deve ser sempre tecnicamente possível.”⁵¹¹

1.2. *Suspender o conteúdo*: consiste na remoção temporária do conteúdo ou em uma suspensão definitiva – o mecanismo, aparentemente, é o mesmo do item anterior, sendo, entretanto, limitado pelo tempo e, a posteriori, podendo retornar-se com o teor da postagem ou removê-la para sempre. A exemplo do caso “*Trump Ban*”, no qual, antes da remoção do conteúdo, o *Facebook* suspendeu temporariamente a postagem.⁵¹²

⁵⁰⁸ CELESTE, Edoardo. Digital punishment: social media exclusion and the constitutionalising role of national courts, *International Review of Law, Computers & Technology*. v. 35, n. 2, p. 162-184, 2021. Disponível em: DOI: 10.1080/13600869.2021.1885106. Acesso em: 31/10/2021.

⁵⁰⁹ Celeste, cita exemplos, em que o Facebook removeu contas de usuários por ter postado comentários que, supostamente, infringiram os termos de uso, sobretudo, sobre a proibição de discurso de ódio: “Facebook considerou o discurso de ódio não apenas a postagem de um usuário invocando o uso de canhões de água e o exército contra requerentes de asilo (LG Frankfurt/Main, 10.09.2018 - 2-03 O 310/ 18 2018), mas também posts em que os usuários citaram uma declaração do primeiro-ministro húngaro Viktor Orbán definindo refugiados sírios como ‘invasores’ muçulmanos (OLG München, 17.07.2018 - 18 W 858/18, sd), ou classificando um jornal como ‘pseudo-esquerdino’, ‘belicista de primeira classe’ e ‘trapo denegridor’ (LG Frankfurt/Main, 14.05.2018 - 2-03 O 182/18 2018)”. Tradução livre do original: “[...] Facebook considered hate speech not only a user’s post invoking the use of water cannons and the army against asylum seekers (LG Frankfurt/Main, 10.09.2018 - 2-03 O 310/18 2018), but also posts in which users quoted a statement by the Hungarian Prime Minister Viktor Orbán defining Syrian refugees as Muslim ‘invaders’ (OLG München, 17.07.2018 - 18 W 858/18, n.d.), or classifying a newspaper as ‘pseudo-lefty’, ‘first-class warmonger’ and ‘denigratory rag’ (LG Frankfurt/Main, 14.05.2018 - 2-03 O 182/18 2018).” (CELESTE, Edoardo. Digital punishment: social media exclusion and the constitutionalising role of national courts, *International Review of Law, Computers & Technology*. v. 35, n. 2, p. 162-184, 2021. p.166. Disponível em: DOI: 10.1080/13600869.2021.1885106. Acesso em: 31/10/2021).

⁵¹⁰ DE GREGORIO, Giovanni. Democratising online content moderation: A constitutional framework. *Computer Law & Security Review*. v. 36, p. 1-28, 2020. p. 27. Disponível em: <https://doi.org/10.1016/j.clsr.2019.105374>. Acesso em: 18/07/2022.

⁵¹¹ Tradução livre do original: “[...] it is necessary that the content previously removed is still available in case online platforms review their first decision. Therefore, the reinstatement of content should always be technically possible.” (DE GREGORIO, Giovanni. Democratising online content moderation: A constitutional framework. *Computer Law & Security Review*. v. 36, p. 1-28, 2020. p. 27. Disponível em: <https://doi.org/10.1016/j.clsr.2019.105374>. Acesso em: 18/07/2022.

⁵¹² UNITED STATES. **Facebook Oversight Board**. Case Decision 2021-001-FB-FBR, OVERSIGHT BD. 9-10. May 5, 2021. Disponível em: <https://www.oversightboard.com/sr/decision/2021/001/pdf-english>. Acesso em: 01/07/2022.

1.3. *Realocar conteúdo*: nesse caso, o conteúdo permanece *online* na plataforma, contudo, é deslocado para um novo “endereço” para um novo URL. Nesse deslocamento há exclusão e reinserção do conteúdo para outro endereço, redefinindo a quantidade de visualizações, além de remover os comentários existentes até o instante da realocação.

1.4. *Editar ou suprimir o conteúdo*: com esse recurso é possível, em vez de remover uma postagem inteira, a plataforma editar ou suprimir partes consideradas violadoras das normas do termo de adesão.

1.5. *Aviso intersticial*: nessa opção, a plataforma insere um aviso antes de o usuário acessar o conteúdo da postagem. Exemplo, telas de aviso do *Facebook*, em casos de nudez adulta e atividades sexuais:

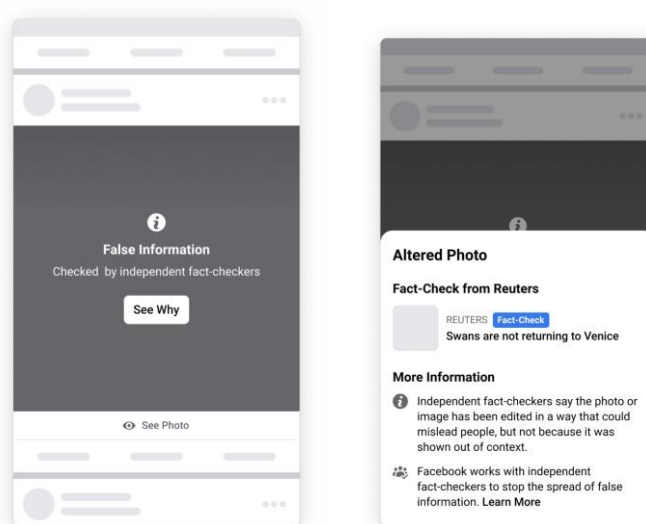


Figura 3: Tela 1 – “Telas de aviso no contexto: Cobrimos determinados conteúdos no Feed de Notícias e em outras plataformas, para que as pessoas possam escolher se querem vê-los.” Tela 2 – “Mais informações: Neste exemplo, damos mais contexto sobre o motivo de termos coberto a foto com informações de verificadores de fatos independentes.”⁵¹³

Nessa mesma linha, tem (1.6) *adicionar legenda de aviso*, (1.7) *inserir contra-fala* (*counterspeech*), e (8) *desabilitar comentários*. Como exemplo de contra-fala, o *Twitter* desenvolveu durante a pandemia do COVID-19, inserindo um *prompt* de pesquisa na barra de informações, o qual chamou de “*Know the Facts*”, para o público encontrar informações claras e confiáveis sobre imunização e saúde vacinal, conforme se verifica na figura abaixo:⁵¹⁴

⁵¹³ META. Padrões da Comunidade do Facebook. **Facebook.com** 2022. Disponível em: <https://transparency.fb.com/pt-br/policies/community-standards/adult-nudity-sexual-activity/>. Acesso em: 15/07/2022.

⁵¹⁴ HARVEY, Del. Helping you find reliable public health information on Twitter. **Blog.Twitter**. 2019. Disponível em: https://blog.twitter.com/en_us/topics/company/2019/helping-you-find-reliable-public-health-information-on-twitter. Acesso em: 18/07/2022.



Figura 4. Twitter insere *prompt* de pesquisa sobre vacinas contra o COVID-19.

2. *Regulamento de Conta*: possui 6 recursos em face da conta de um usuário.

2.1. *Encerrar conta*: representa um remédio padrão, configurando-se na simples remoção de contas permanentemente;

2.2. *Suspender conta*: impede, temporariamente, – que pode ir de minutos até em definitivo – que os usuários acessem suas contas;

2.3. *Suspender direitos de postagem*: esse remédio susta a capacidade do usuário em realizar postagens, mesmo permanecendo a conta *online*. O *Facebook*, utiliza esse recurso com a denominação *restrição de usuários* – a qual também é aplicada para grupos e páginas⁵¹⁵ – nota-se, para a maioria das violações, o primeiro passo ser uma advertência com um aviso. E, caso ocorram novas violações, somando duas ou mais advertências, serão aplicadas outras restrições, vejamos a previsão:

Na maioria das violações no Facebook, as advertências causam as seguintes restrições:

Uma advertência: aviso e nenhuma restrição.

Duas advertências: restrição de um dia à criação de conteúdo, como publicar, comentar, usar o Facebook Live ou criar uma Página.

Três advertências: restrição de três dias à criação de conteúdo.

Quatro advertências: restrição de sete dias à criação de conteúdo.

Cinco advertências ou mais: restrição de 30 dias à criação de conteúdo.⁵¹⁶

2.4. *Remover emblemas de credibilidade*: as plataformas *online* em casos de violação das regras, podem retirar o símbolo ou ícone da conta do usuário, que atesta que o perfil era verificado e confiável, diminuindo a credibilidade daquele indivíduo.

⁵¹⁵ Quando essa aplicação, assim dispõe: “Páginas e grupos no Facebook que violam nossas políticas de forma recorrente podem ser removidos das recomendações e ter sua distribuição reduzida. As páginas também podem sofrer restrições nos recursos de monetização e os grupos podem ser obrigados a ter as publicações aprovadas pelo administrador. Se você gerencia uma Página ou grupo, é possível encontrar informações sobre violações e restrições em Qualidade da Página ou Qualidade do Grupo.” (META. Tomada de medidas: Restrição de usuários. **Transparecny Center**. 2022. Disponível em: <https://transparency.fb.com/pt-br/enforcement/taking-action/restricting-accounts/>. Acesso em: 14/07/2022)

⁵¹⁶ META. Tomada de medidas: Restrição de usuários. **Transparency Center**. 2022. Disponível em: <https://transparency.fb.com/pt-br/enforcement/taking-action/restricting-accounts/>. Acesso em: 14/07/2022

2.5. *Reduzir os níveis de serviço*: esse remédio consiste em reduzir a oferta de serviços da conta, como limitar o número de vezes que uma postagem pode ser acessada, lida ou visualizada;

2.6. *Envergonhar (Shaming)*: nesse caso a plataformas ou serviço de *internet*, utiliza uma forma de atrair a atenção pública para o comportamento considerado violador do usuários, como forma de inibir uma nova conduta deste.

Esse remédio deve ser bem elaborado pelas plataformas *online*, a fim de evitar sua aplicação danosa ao usuário “envergonhado”, passando este a ser considerado como vítima. Nesse sentido, Laidlaw resume: “às vezes, os envergonhados cometeram uma transgressão social e as turbas da *internet* usam sanções de vergonha para impor as normas da comunidade *online* [...]. Outras vezes, os envergonhados não necessariamente cometeram uma transgressão social, mas são alvos mesmo assim”.⁵¹⁷ Entende, ainda, que a vergonha pode impor regras de civilidade em comunidades *online*, podendo ser considerada como uma força facilitadora para mudanças positivas, afirmando “*naming and shaming* é uma ferramenta regulatória central para lidar com os abusos dos direitos humanos.”⁵¹⁸

Por outro lado, é importante não confundir esse remédio com o problema do *shaming online*, o qual, segundo Kate Klonick, surgiu em decorrência de um ambiente digital propício cujas restrições naturais à vergonha, como o acesso à *Internet* de baixo custo, anônimo, instantâneo e onipresente sejam demonstradas publicamente. Em face disto, discute-se a necessidade de intervenção do Estado para conter a vergonha *online* e o assédio cibernético. Surgindo, nesse debate, a viabilidade das soluções legais, normativas, privadas e estatais para controlar o *shaming online*.⁵¹⁹ Dentre elas, a “aplicação de normas *online* e remédios privados para informar o debate em torno da intervenção estatal”.⁵²⁰

⁵¹⁷ LAIDLAW, Emily B. Online Shaming and the Right to Privacy. **Computer Science Laws**. v. 6, n. 3, p. 1-26, 2017. p. 2. Disponível em: <https://doi.org/10.3390/laws6010003> ou https://pdfs.semanticscholar.org/d08d/92023799353d3fd150896c731c0d8166ebc8.pdf?_ga=2.237340480.1368221140.1660520780-227723297.1658242439. Acesso em: 20/07/2022.

⁵¹⁸ LAIDLAW, Emily B. Online Shaming and the Right to Privacy. **Computer Science Laws**. v. 6, n. 3, p. 1-26, 2017. p. 2. Disponível em: <https://doi.org/10.3390/laws6010003> ou https://pdfs.semanticscholar.org/d08d/92023799353d3fd150896c731c0d8166ebc8.pdf?_ga=2.237340480.1368221140.1660520780-227723297.1658242439. Acesso em: 20/07/2022.

⁵¹⁹ KLONICK, Kate. Re-Shaming the Debate: Social Norms, Shame, and Regulation in a Internet Age. **Maryland Law Review**. v. 75. n. 4, p. 1029-1065, 2016. p. 1029. Disponível em: <https://ssrn.com/abstract=2638693> ou <http://dx.doi.org/10.2139/ssrn.2638693>. Acesso em: 20/07/2022.

⁵²⁰ Tradução livre do original: “[...] online norm enforcement, and private remedies to inform the debate around state intervention.” (KLONICK, Kate. Re-Shaming the Debate: Social Norms, Shame, and Regulation in a Internet Age. **Maryland Law Review**. v. 75. n. 4, p. 1029-1065, 2016. p. 1030. Disponível em: <https://ssrn.com/abstract=2638693> ou <http://dx.doi.org/10.2139/ssrn.2638693>. Acesso em: 20/07/2022).

3. *Restrições de visibilidade*: nessa categoria os remédios atuam para diminuir o alcance do conteúdo postado pelo usuário, seja de forma parcial ou total. É a categoria que mais possuem remédios.

3.1. *Shadowban*: consiste, essencialmente, quando uma plataforma de mídia social oculta o conteúdo de alguém, pelo algoritmo com táticas como torná-los invisíveis nas hashtags, proibir curtidas/comentários ou censurar, continuamente, seu conteúdo.⁵²¹ Nessa mesma senda, West afirma que *shadowbanning* ocorre quando o conteúdo de uma pessoa ou a visibilidade da conta é muito reduzida, ou se torna invisível para outros usuários sem realmente ser removido completamente.⁵²² Portanto, a conta do usuário permanece ativa, contudo, apenas o titular da conta tem acesso ao conteúdo.⁵²³ Suzor, West, Quodling, e York, acrescentam que, muitas vezes, as plataformas adotam os chamados “*shadowbans*”, quando decidem ocultar ou tirar a prioridade de conteúdo sem informar ao usuário. Reconhecem, também, os *shadowbans* com utilidade para lidar com *spammers* que criam *bots* para postar conteúdo. Contudo, os pesquisadores apontam, devido à sua própria natureza, os *shadowbans*, com possibilidade de criar incertezas e causar nos usuários a sensação de terem sido censurados de forma oculta, sem notificação.⁵²⁴

Como forma de evitar tal sentimento, Suzor, West, Quodling, e York recomendam às plataformas *online* uma transparência significativa na moderação de conteúdo, e “sugerimos que as plataformas garantam que os usuários sejam informados quando seu conteúdo for removido ou tornado invisível para públicos específicos.”⁵²⁵ Além disso, essas “notificações devem estar permanentemente disponíveis para o usuário de alguma forma, e não apenas enviadas como notificações no aplicativo potencialmente efêmeras.”⁵²⁶

⁵²¹ HORNE, Julia Van. Shadowbanning is a Thing — and It’s Hurting Trans and Disabled Advocates. **Salty**. 2020. Disponível em: <https://saltyworld.net/shadowbanning-is-a-thing-and-its-hurting-trans-and-disabled-advocates>. Acesso em: 18/07/2022.

⁵²² WEST, Sarah Myers. Censored, suspended, shadowbanned: User interpretations of content moderation on social media platforms. **New Media & Society**. v. 20, n. 11, p. 4366-4383, 2018. Disponível em: <https://doi.org/10.1177/1461444818773059>. Acesso em: 19/07/2022.

⁵²³ Conferir também: HARRIS, Sukanya; JACKSON, Melody Moore; ZEAGLER, Clint. **Content Censorship on Social Media and the Impact on the Content Creator Communities**. 2022. Disponível em: https://sukanyaharris.com/wp-content/uploads/2022/05/SukanyaHarris_TermPaper.pdf. Acesso em: 15/07/2022.

⁵²⁴ SUZOR, Nicolas P.; WEST, Sarah Myers; QUODLING, Andrew; YORK, Jillian. What Do We Mean When We Talk About Transparency? Toward Meaningful Transparency in Commercial Content Moderation. **International Journal of Communication**. v. 13, p. 1526–1543. 2019. p. 1531. Disponível em: <https://ijoc.org/index.php/ijoc/article/view/9736>. Acesso em: 19/07/2022.

⁵²⁵ Tradução livre do original: “[...] we suggest that platforms ensure that users are informed when their content is removed or made invisible to particular audiences.” (SUZOR, Nicolas P.; WEST, Sarah Myers; QUODLING, Andrew; YORK, Jillian. What Do We Mean When We Talk About Transparency? Toward Meaningful Transparency in Commercial Content Moderation. **International Journal of Communication**. v. 13, p. 1526–1543. 2019. p. 1532. Disponível em: <https://ijoc.org/index.php/ijoc/article/view/9736>. Acesso em: 19/07/2022).

⁵²⁶ Tradução livre do original: “[...] notifications should be permanently available to the user in some form, and not just sent as potentially ephemeral in-app notifications.” (SUZOR, Nicolas P.; WEST, Sarah Myers;

3.2. *Remover do índice de pesquisa externa*: com o crescimento exponencial do conteúdo na *internet*, os motores de busca tornaram-se o ponto de partida para a maioria dos sites. Assim, esse remédio permite que seja fixado uma *tag*⁵²⁷ de não indexar – “*noindex*” em determinada página, excluindo dos índices de pesquisa externos.

3.3 *Links dos autores do Nofollow*: assim como o remédio anterior, nesse o serviço insere uma *tag* “*nofollow*” que determina que os *links*, eventualmente existentes na página, não sejam seguidos.⁵²⁸

3.4. *Remover do índice de pesquisa interna*: consiste na remoção do conteúdo do índice próprio da plataforma ou sítio. Por exemplo, o *Instagram* exclui determinado tipo de postagens de sua pesquisa de *hashtag*.

3.5. *Downgrade da visibilidade da pesquisa interna*: evita-se a remoção do conteúdo do índice interno, optando por reduzir a visibilidade do conteúdo nos resultados das pesquisas. Edward Lee exemplifica esse remédio com um fato ocorrido em outubro de 2020, no qual um controverso artigo do *New York Post*, relatava supostos e-mails de um empresário ucraniano para Hunter Biden, indicando uma suposta reunião com Joe Biden e buscando usar a “influência” de Hunter Biden, a partir do qual o *Facebook* permitiu *links* para o artigo do *New York Post*, mas rebaixou (*downgraded*) sua proeminência nos *feeds* de notícias dos usuários no *Facebook*, aguardando a verificação de fatos.⁵²⁹

3.6. *Sem sugestão automática*: remédio que, simplesmente, não dispõe o conteúdo considerado violador no mecanismo de pesquisa interna de um serviço, do modo “sugestão automática”;

3.7. *Não/promoção interna reduzida*: remédio que reúne característica do *downgrade*, só que nesse caso, a redução da visibilidade dar-se-á em promoções internas, incluindo

QUODLING, Andrew; YORK, Jillian. What Do We Mean When We Talk About Transparency? Toward Meaningful Transparency in Commercial Content Moderation. **International Journal of Communication**. v. 13, p. 1526–1543. 2019. p. 1532. Disponível em: <https://ijoc.org/index.php/ijoc/article/view/9736>. Acesso em: 19/07/2022).

⁵²⁷ “A extração das informações de entrada nas páginas *Web* é feita através da leitura de suas *tags*. *Tags* são estruturas de linguagem de marcação, que podem ser interpretadas por navegadores *Web*, que consistem em breves instruções com uma marca de início e outra de fim.” (MORAIS, Edison A.; AMBRÓSIO, Ana Paula L. Ferramentas de busca na Internet. **Relatório Técnico: Universidade Federal de Goiás**. 2007. p. 8. Disponível em: https://ww2.inf.ufg.br/sites/default/files/uploads/relatorios-tecnicos/RT-INF_002-07.pdf. Acesso em: 15/07/2022).

⁵²⁸ MORAIS, Edison A.; AMBRÓSIO, Ana Paula L. Ferramentas de busca na Internet. **Relatório Técnico: Universidade Federal de Goiás**. 2007. p. 9. Disponível em: https://ww2.inf.ufg.br/sites/default/files/uploads/relatorios-tecnicos/RT-INF_002-07.pdf. Acesso em: 15/07/2022).

⁵²⁹ LEE, Edward. Moderating Content Moderation: A Framework for Nonpartisanship in Online Governance. **American University Law Review**. v. 70, p. 913-1059, 2021. p. 940. Disponível em: http://aulawreview.org/au_law_review/wp-content/uploads/2021/02/Lee.to_.Printer.pdf. Acesso em: 18/07/2022.

recomendações. O *Facebook*, por exemplo, reduz a visibilidade de páginas/contas que exibem conteúdos falsos.⁵³⁰

3.8. *Não/Links de navegação reduzidos*: outro remédio em favor da liberdade de expressão, não removendo o conteúdo somente⁵³¹. Esse recurso reduz a exposição do conteúdo do índice de navegação interna. Podendo fazer um *downrank* do conteúdo, o qual consiste em reduzir o número de vezes que uma postagem aparece nos *feeds* de mídia social de outros usuários. Perez aponta, como exemplo, o *Facebook* rebaixar (*downrank*) “alegações de saúde exageradas ou sensacionais, bem como aqueles que tentam vender produtos ou serviços com base em alegações relacionadas à saúde.”⁵³²

3.9. *Viralidade reduzida*: esse tem por objetivo, como visto em alguns remédios dessa categoria, reduzir a disseminação viral ou a visibilidade;

3.10. *Age-Gate*: consiste, tão somente, em serviço de elaborar uma classificação do conteúdo ou página de acordo com a idade, como restringir páginas apenas para adultos;

3.11. *Exibir conteúdo apenas para leitores conectados*: com isso, restringe a abrangência do conteúdo e permite o acesso somente para usuários previamente registrados, limitando o acesso ao público em geral.⁵³³

4. *Monetária*: essa categoria de remédios são aplicáveis às plataformas ou serviços de *Internet* nos quais são cobrados algum pagamento aos usuários pelo conteúdo, ou de alguma forma detém dinheiro dos usuários.

⁵³⁰ Vejamos como o Facebook trata esse remédio: “Garantir que menos pessoas vejam informações incorretas: quando um verificador de fatos classifica um conteúdo como falso, alterado ou parcialmente falso, ele aparece em uma posição inferior no Feed. Não é exibido no Explorar no Instagram e é apresentado com menos destaque no Feed e nos Stories. Isso reduz significativamente o número de pessoas que o visualizam. Também rejeitamos anúncios com conteúdo avaliado por verificadores de fatos.” (META. Sobre a verificação de fatos no Facebook. **Central de Ajuda do Meta Business**. 2021. Disponível em: <https://www.facebook.com/business/help/2593586717571940>. Acesso em: 15/07/2022).

⁵³¹ Nesse sentido, Saltz, Barari, Leibowicz e Wardle, afirmam: “Além disso, nem todo tipo de resposta da plataforma online à desinformação ou intervenção é igualmente grave. Alguns, como a remoção, impedem totalmente o acesso, levantando preocupações sobre a liberdade de expressão, enquanto o **downranking** dificulta o acesso ao reduzir o alcance.” (Grifo nosso) Tradução livre do original: “Beyond this, not every type of online platform response to misinformation, or intervention, is equally severe. Some, like removal, prevent access entirely, raising free speech concerns, whereas downranking makes access more difficult by reducing reach.” (SALTZ, Emily; BARARI, Soubhik; LEIBOWICZ, Claire; WARDLE, Claire. Misinformation interventions are common, divisive, and poorly understood. **Harvard Kennedy School Misinformation Review**. v. 2, n. 5, p. 1-25, 2021. p. 2. Disponível em: https://misinforeview.hks.harvard.edu/wp-content/uploads/2021/10/saltz_misinformation_interventions_20211027.pdf. Acesso em: 18/07/2022).

⁵³² Tradução livre do original: “[...] exaggerated or sensational health claims, as well as those trying to sell products or services based on health-related claims.” (PEREZ, Sarah. Facebook news feed changes downrank misleading health info and dangerous ‘cures’. **TechCrunch**. 2019. Disponível em: <https://techcrunch.com/2019/07/02/facebook-news-feed-changes-downrank-misleading-health-info-and-dangerous-cures/>. Acesso em: 18/07/2022).

⁵³³ GOLDMAN, Eric. Content Moderation Remedies. **Michigan Technology Law Review**. v. 28, n. 1. p. 1-59, 2021. p. 35. Disponível em: <https://repository.law.umich.edu/mtlr/vol28/iss1/2>. Acesso em: 02/11/2021.

4.1. *Perda de ganhos acumulados*: o serviço pode reter os ganhos acumulados em caso de violação. Por exemplo, o *Google* pode reter o lucro acumulado de editores que venham a violar as regras do *Google AdSense*.⁵³⁴

4.2. *Encerrar Ganhos Futuros (por item ou conta)*: consiste na plataforma ou serviço encerrar o pagamento de futuros lucros, também conhecido como desmonetização. A exemplo do *YouTube* que conta com algoritmos na moderação de conteúdo e “segue uma lógica de justiça retributiva para punir os *YouTubers* condenados por meio da desmonetização, uma penalidade que limita ou os priva de anúncios, reduzindo sua receita futura de anúncios.”⁵³⁵

4.3. *Suspender ganhos futuros (por item ou conta)*: diferentemente do anterior, este não encerra, definitivamente, os ganhos; a plataforma ou serviço apenas suspende, por determinado tempo, essa capacidade de obter lucros.

4.4. *Autor da multa/Impor danos liquidados*: a plataforma ou serviço pode reter parte ou totalidade de dinheiro como forma de aplicar uma penalidade – assim, não difere do remédio anterior –. Ou, de mesmo modo, impor uma “multa” ao usuário de acordo com os Termos de Serviço.⁵³⁶

5. *Outros*: nessa categoria, são reunidos os remédios que não se enquadram em nenhuma das anteriores.

5.1. *Educar os usuários*: esse remédio é essencial para que não ocorra, ou evitar, violações de regras, devendo as plataformas *online* e serviços de *Internet* ensinar ao usuário sobre as regras e normas do serviço. Nesse sentido, os *hosts* devem educar os indivíduos sobre suas políticas de moderação de conteúdo e o meio pelos quais são aplicadas. Kamara, Kodel, Llansó, Nojeim, Qin, Thakur e Vogus, em importante estudo sobre a moderação de conteúdo, aduzem que a fase de educação pode assumir várias formas, sendo importante

⁵³⁴ “O Google AdSense é uma plataforma de publicidade que utiliza várias informações sobre o usuário que visualiza os anúncios, para mostrar a ele os mais relevantes e assim maximizar sua própria renda. As informações mais comuns sobre o usuário são sua geolocalização, os idiomas que ele fala, seus interesses e o conteúdo do site que ele está navegando no momento.” (DESNICA, Marin. SAYADCHI, Marie; KHAWJA, Umair; SZABÓ, Peter. **Google AdSense: User Modeling and Recommender Systems – Case exercise**, 2014. Disponível em: <http://htk.tlu.ee/recommendersystems/images/5/5b/Google-adsense-v1.1.pdf>. Acesso em: 18/07/2022).

⁵³⁵ Tradução livre do original: “[...] follows a retributive justice logic to punish convicted YouTubers through demonetization, a penalty that limits or deprives them of advertisements (ads), reducing their future ad income.” (MA, Renkai; KOU, Yubo. “How advertiser-friendly is my video?”: Youtuber’s Socioeconomic Interactions with Algorithmic Content Moderation. **Proceedings of the ACM on Human-Computer Interaction**, v. 5, n. CSCW2, p. 1-25, 2021. p. 1. Disponível em: https://www.researchgate.net/profile/Renkai-Ma-2/publication/355078337_How_advertiser-friendly_is_my_video_Youtuber's_Socioeconomic_Interactions_with_Algorithmic_Content_Moderation/links/615c761dc04f5909fd8560f9/How-advertiser-friendly-is-my-video-Youtubers-Socioeconomic-Interactions-with-Algorithmic-Content-Moderation.pdf. ou <https://doi.org/10.1145/3479573>. Acesso em: 18/07/2022)

⁵³⁶ GOLDMAN, Eric. Content Moderation Remedies. **Michigan Technology Law Review**. v. 28, n. 1. p. 1-59, 2021. p. 36. Disponível em: <https://repository.law.umich.edu/mtlr/vol28/iss1/2>. Acesso em: 02/11/2021.

assumir o compromisso de informar sobre a mais básica, que são “os termos de serviço do serviço, as diretrizes da comunidade e outras informações voltadas para o usuário sobre as políticas do site.”⁵³⁷

De igual modo, os moderadores também podem orientar os usuários sobre qual conteúdo é permissível ou não. Como sublinha Grimmelmann, os moderadores podem influenciar nas normas “elogiando o bom comportamento e criticando o mau” ou fornecendo explicações aos usuários quando seu conteúdo é moderado.⁵³⁸ Além disso, os moderadores podem incluir nas notificações enviadas aos usuários informações sobre a tomada da decisão e o motivo pelo qual foi tomada, além disso, é possível explicar o porquê de uma apelação feita por eles ter sido negada. A educação é um componente crítico da moderação de conteúdo, especialmente, em resposta a violações de boa-fé das políticas de moderação de conteúdo.⁵³⁹ É de suma importância que as plataformas e serviços de *Internet*, instrua seus usuários, para que possam ser capazes de entender quais tipos de conteúdo são ou não permitidos, para que o processo de moderação de conteúdo funcione de forma eficaz e transparente.⁵⁴⁰

Para arrematar, demonstrando a importância da educação do usuário como meio de transformar a moderação de conteúdo, minimizando danos para as plataformas e maximizando a liberdade de expressão e privacidade dos usuários, Jhaver, Bruckman e Gilbert concluem que:

Nossas descobertas mostram que o fornecimento de explicações de remoção está associado a uma redução nas remoções futuras, sugerindo que adotar uma abordagem educacional, em vez de punitiva, à moderação de conteúdo pode melhorar os resultados da comunidade. Nossa análise também indica que o uso de ferramentas automatizadas para fornecer explicações de remoção é uma abordagem

⁵³⁷ Tradução livre do original: “[...] the service’s terms of service, Community Guidelines, and other user-facing information about the site’s policies.” (KAMARA, Seny; KNODEL, Mallory; LLANSÓ, Emma; NOJEIM, Greg; QIN, Lucy; THAKUR, Dhanaraj; VOGUS, Caitlin. *Outside Looking In: Approaches to Content Moderation in End-to-End Encrypted Systems*. **Center for Democracy & Technology**. 2022. p. 11. Disponível em: <https://arxiv.org/ftp/arxiv/papers/2202/2202.04617.pdf> ou <https://doi.org/10.48550/arXiv.2202.04617>. Acesso em: 18/07/2022).

⁵³⁸ Tradução livre do original: “[...] by praising good behavior and criticizing bad.” (GRIMMELMANN, James. *The Virtues of Moderation*. **The Yale Journal of Law & Technology**. v. 17, n. 1, p. 42-109, 2015. p. 62. Disponível em: <https://core.ac.uk/download/pdf/72836082.pdf> ou <https://digitalcommons.law.yale.edu/yjolt/vol17/iss1/2>. Acesso em: 10/01/2022).

⁵³⁹ KAMARA, Seny; KNODEL, Mallory; LLANSÓ, Emma; NOJEIM, Greg; QIN, Lucy; THAKUR, Dhanaraj; VOGUS, Caitlin. *Outside Looking In: Approaches to Content Moderation in End-to-End Encrypted Systems*. **Center for Democracy & Technology**. 2022. p. 11. Disponível em: <https://arxiv.org/ftp/arxiv/papers/2202/2202.04617.pdf> ou <https://doi.org/10.48550/arXiv.2202.04617>. Acesso em: 18/07/2022.

⁵⁴⁰ JHAVER, Shagun; BRUCKMAN, Amy; GILBERT, Eric. Does Transparency in Moderation Really Matter?: User Behavior After Content Removal Explanations on Reddit. **Proceedings of the ACM on Human-Computer Interaction**. v. 3, n. CSCW, Article 150. ACM, New York, NY. 2019. p. 150:2. Disponível em: <https://doi.org/10.1145/3359252> ou <https://dl.acm.org/doi/pdf/10.1145/3359252>. Acesso em: 15/07/2022

promissora para projetar transparência sem aumentar indevidamente a carga de trabalho dos moderadores.⁵⁴¹

5.2. *Atribuir avisos/avisos*: na mesma linha de pensamento apresentada no subitem anterior, informar ao usuário que seu comportamento viola as regras e normas.

5.3. *Saída/desmascaramento*: ação a partir da qual a plataforma *online* pode revelar a verdadeira identidade de um usuário com uso de pseudônimo, o que pode resultar no *Shamming* (visto no subitem 2.6).

5.4. *Reportar à aplicação da lei*: o serviço poderá relatar as violações que também configurem transgressão da lei para que sejam processadas e julgadas.

5.5. *Coloque o usuário/conteúdo na lista de bloqueio de toda a indústria*: as plataformas *online* podem compartilhar entre si a listagem de usuários que violaram as suas regras, e, assim, construírem “listas de bloqueios” as quais abrangem todo o setor para usuários ou conteúdo específicos.⁵⁴² Por exemplo, a lista de bloqueio do *Twitter*, como explicam Jhaver, Ghoshal, Bruckman e Gilbert, permitem que “usuários bloqueiem rapidamente todas as contas em uma lista de contas dignas de bloqueio geradas por algoritmos ou com curadoria da comunidade”.⁵⁴³ Ademais, esse mecanismo de moderação descentralizado, como listas de bloqueio anti-assédio, “tira um pouco da pressão dos moderadores centrais do *Twitter* para que eles não precisem ser tão rigorosos em sua moderação.”⁵⁴⁴

5.6. *Serviço comunitário*: prestação de serviço à comunidade como forma de reparar alguma situação de violação da regra.

⁵⁴¹ Tradução livre do original: “Our findings show that provision of removal explanations is associated with a reduction in future removals, suggesting that taking an educational, rather than a punitive, approach to content moderation can improve community outcomes. Our analysis also indicates that using automated tools to provide removal explanations is a promising approach to design for transparency without unduly increasing the work load of moderators.” (JHAVER, Shagun; BRUCKMAN, Amy; GILBERT, Eric. Does Transparency in Moderation Really Matter?: User Behavior After Content Removal Explanations on Reddit. **Proceedings of the ACM on Human-Computer Interaction**. v. 3, n. CSCW, Article 150. ACM, New York, NY. 2019. p. 150:24. Disponível em: <https://doi.org/10.1145/3359252> ou <https://dl.acm.org/doi/pdf/10.1145/3359252>. Acesso em: 15/07/2022).

⁵⁴² GOLDMAN, Eric. Content Moderation Remedies. **Michigan Technology Law Review**. v. 28, n. 1. p. 1-59, 2021. p. 39. Disponível em: <https://repository.law.umich.edu/mtlr/vol28/iss1/2>. Acesso em: 02/11/2021.

⁵⁴³ Tradução livre do original: “[...] users to quickly block all accounts on a community-curated or algorithmically generated list of block-worthy accounts.” (JHAVER, Shagun; GHOSHAL, Sucheta; BRUCKMAN, Amy; GILBERT, Eric. Online Harassment and Content Moderation: The Case of Blocklists. **ACM Transactions on Computer-Human Interaction**. v. 25, n. 2, Article 12, 2018. p. 12:3. Disponível em: <https://shagunjhaver.com/research/articles/jhaver-2018-blocklists/jhaver-2018-blocklists.pdf>. Acesso em: 18/07/2022).

⁵⁴⁴ Tradução livre do original: “[...] takes some pressure off the central Twitter moderators so that they don’t have to be as strict in their moderation.” (JHAVER, Shagun; GHOSHAL, Sucheta; BRUCKMAN, Amy; GILBERT, Eric. Online Harassment and Content Moderation: The Case of Blocklists. **ACM Transactions on Computer-Human Interaction**. v. 25, n. 2, Article 12, 2018. p. 12:3. Disponível em: <https://shagunjhaver.com/research/articles/jhaver-2018-blocklists/jhaver-2018-blocklists.pdf>. Acesso em: 18/07/2022).

5.7. *Justiça Restaurativa/Desculpas*: a fim de dirimir as consequências danosas à sociedade, o indivíduo que violou as regras poderá pedir desculpas aos usuários afetados, ou então participar de algum processo de justiça restaurativo com o objetivo de debater, juntamente à comunidade a afetação das infrações às normas.

Apresentamos os principais remédios, os quais visam corrigir violações, evitando – ao máximo – a remoção definitiva de conteúdo ou contas. Como é possível auferir da construção lógica dessa pesquisa, os remédios podem ser aplicados tanto de forma autônoma, quanto de forma combinada, sempre prestigiando a manutenção do conteúdo/conta.

Imperativo mencionar os remédios alternativos como a demonstração do uso eficaz das ferramentas pelos moderadores. É cediço que nenhum remédio terá eficácia para a maioria das correções, até porque, são inúmeras as formas e, a cada dia, surgem novas tipificações de infração. Destarte, muitos remédios funcionam a um determinado tipo de plataformas *online*, podendo ser ineficientes a outras.

Importante a reflexão de Douek, que dada a vastidão de discursos que a *internet* possibilitou, a qual deverá ser governada pela moderação de conteúdo, tomou uma proporção inimaginável. Com essa grande escala de produção de conteúdo, “uma decisão ‘correta’ em cada caso é praticamente impossível, nesse volume, falhas individuais não podem ser consideradas falhas do sistema.”⁵⁴⁵

Eric Goldman, assinala os seguintes fatores para escolha a ser feita pelos reguladores ou moderadores dentre os remédios alternativos, qual melhor se adequaria ao caso, à plataforma, ao tipo de moderação: (i) *Gravidade da violação da regra*; (ii) *Confiança de que uma violação de regra realmente ocorreu*; (iii) *Escalabilidade e consistência*; (iv) *A capacidade da comunidade de autocorreção*; (v) *Como os remédios afetam outras pessoas*; (vi) *Reter o envolvimento do usuário enquanto restringe as violações e reincidência*; (vii) *Sanções paralelas*.⁵⁴⁶

⁵⁴⁵ Tradução livre do original: “[...] a “correct” decision in each case (assuming, for the sake of argument, there is such a thing) is practically impossible at this volume, individual failures cannot be considered system failures.” (DOUEK, Evelyn, *The Siren Call of Content Moderation Formalism* (January 10, 2022). In: BOLLIGER, Lee; GEOFFREY, Stone (Eds.). **New Technologies of Communication and the First Amendment: The Internet, Social Media and Censorship**. Oxford Press, forthcoming 2022. p. 11. Disponível em: <https://ssrn.com/abstract=4005314> ou <http://dx.doi.org/10.2139/ssrn.4005314>. Acesso em: 20/07/2022).

⁵⁴⁶ Tradução livre do original: “severity of the rule violation; confidence that a rule violation actually occurred; scalability and consistency; the community’s ability to self-correct; how the remedies impact others; retaining user engagement while curbing violations and recidivism; parallel sanctions.” (GOLDMAN, Eric. *Content Moderation Remedies*. **Michigan Technology Law Review**. v. 28, n. 1. p. 1-59, 2021. p. 41. Disponível em: <https://repository.law.umich.edu/mtlr/vol28/iss1/2>. Acesso em: 02/11/2021).

Logicamente, não é uma lista exaustiva, são apenas vetores de aplicação dos remédios, devendo levar em consideração a:

(i) *gravidade da violação da regra*, é dizer, deve ser proporcional à violação, desafiando remédios mais severos de acordo com a gravidade da postura do usuário. Com esse pensamento, Land e Hamilton abordam que “embora os remédios para danos aos direitos humanos sejam direcionados a uma variedade de objetivos, um princípio central da lei de reparação é a proporcionalidade. Qualquer remédio escolhido deve ser proporcional à gravidade e dano da violação”.⁵⁴⁷

(ii) *Confiança de que uma violação de regra realmente ocorreu*: nem sempre qualquer conteúdo produzido pelo usuário terá uma gravidade explícita à regra da plataforma, o que demandará uma análise do contexto social e local, para, por exemplo, definir se o conteúdo é ou não difamatório. Na prática, nota-se a maioria das plataformas *online* optando por remover o conteúdo – seja por meio de forma automatizada, por algoritmos, ou por meio de denúncias de usuários que se sintam ofendidos, sem que de fato tenha uma apuração legítima e satisfatória quanto à violação e da extensão de sua gravidade.

Assim, as soluções de não remoção tendem a dar alternativas às plataformas para as situações em que não tiverem plena convicção quanto à violação. Desse modo, nos casos que não sejam tão evidentes, aplicar-se-iam remédios alternativos e menos severos.⁵⁴⁸ Ademais, essas medidas evitam as remoções de “falso positivo” – consistindo na exclusão de conteúdo que não deveria ser removido, mas, por falta de melhor análise, é considerado como conteúdo violador⁵⁴⁹ – tal como a remoção indevida de conteúdo benéfico sobre máscaras destinadas a profissionais de saúde, conforme apontado no início desse capítulo.

(iii) *Escalabilidade e consistência*: no processo de moderação de conteúdo, a atuação em escala e sua consistência são fundamentais, pois, devido ao grande volume de conteúdo produzido nas plataformas, é inviável uma moderação individualizada. Como assinala Gillespie, a “quantidade, velocidade e variedade de conteúdo é estratosférica; os usuários

⁵⁴⁷ Tradução livre do original: “Although remedies for human rights harms are directed to a variety of objectives, a central principle of the law on remedy is proportionality. Any remedy chosen must be proportional to the gravity and harm of the violation”. (LAND, Molly K.; HAMILTON, Rebeca J. *Beyond takedown: Expanding the tool kit for responding to online hate*. In: DOJČINOVIĆ, Predrag (Ed.). **Propaganda and International Criminal Law: From cognition to Criminality**. 1 ed. London: Routledge, 2019. p. 7. Disponível em: <https://doi.org/10.4324/978042944369>. Acesso em: 01/12/2021).

⁵⁴⁸ GOLDMAN, Eric. Content Moderation Remedies. **Michigan Technology Law Review**. v. 28, n. 1. p. 1-59, 2021. p. 43. Disponível em: <https://repository.law.umich.edu/mtlr/vol28/iss1/2>. Acesso em: 02/11/2021.

⁵⁴⁹ Para uma melhor análise, conferir: DOUEK, Evelyn. Governing online speech: from “posts-as-trumps” to proportionality and probability. **Columbia Law Review**. v. 121, n. 3, p. 759-837, 2021. Disponível em: https://www.columbialawreview.org/wp-content/uploads/2021/04/Douek-Governing_Online_Speech-from_Posts_As-Trumps_To_Proportionality_And_Probability.pdf. Acesso em: 03/11/2021. p. 808-seguintes.

estão ligados menos por laços de comunidade e mais por algoritmos de recomendação e gráficos sociais”.⁵⁵⁰

Além disso, aumentariam os riscos de ocorrência de erros, aumento nos custos, a escala, portanto, é fundamental na moderação de conteúdo.⁵⁵¹ Nesse sentido, as plataformas têm adotado a moderação algorítmica como maneira de identificar violações – inclusive, como já previa Lawrence Lessig, há mais de uma década, o *código* tornar-se-ia um regulador cada vez mais importante.⁵⁵²

Como os sistemas de moderação algorítmica⁵⁵³ estão sendo mais implantados e utilizados, em escala, pelas principais plataformas de conteúdo tais como: *Facebook*, *YouTube* e *Twitter* para controle de conteúdo. Tem-se percebido, por outro lado, conforme pregoam Gorwa, Binns e Katzenbach, o aumento de muitos problemas com a política de conteúdo praticada pelas plataformas, por isso, a moderação automatizada tem representado uma ameaça ao:

(a) aumentar ainda mais a opacidade, tornando um conjunto de práticas notoriamente não transparentes ainda mais difícil de entender ou auditar, (b) complicar ainda mais questões pendentes de equidade e justiça em sistemas sociotécnicos de grande escala e (c) obscurecer novamente a natureza fundamentalmente política de decisões de fala sendo executadas em escala.⁵⁵⁴

⁵⁵⁰ Tradução livre do original: “The quantity, velocity, and variety of content is stratospheric; users are linked less by the bonds of community and more by recommendation algorithms and social graphs; the consequences of online harms now extend beyond the platform on which they occur; [...]” (GILLESPIE, Tarleton. Content Moderation, AI, and the question of scale. **Big Data & Society**. 2020. p. 1. Disponível em: <https://journals.sagepub.com/doi/full/10.1177/2053951720943234>. ou <https://10.1177/2053951720943234>. Acesso em: 10/07/2022).

⁵⁵¹ Douek reforça que: “escala do conteúdo online fará com que o papel das plataformas privadas como atores da linha de frente na moderação de conteúdo seja uma necessidade prática contínua.” Tradução livre do original: “[...] the scale of online content will make private platforms’ role as the frontline actors in content moderation an ongoing practical necessity.” (DOUEK, Evelyn. Governing online speech: from “posts-as-trumps” to proportionality and probability. **Columbia Law Review**. v. 121, n. 3, p. 759-837, 2021. p. 768. Disponível em: https://www.columbialawreview.org/wp-content/uploads/2021/04/Douek-Governing_Online_Speech-from_Posts_As-Trumps_To_Proportionality_And_Probability.pdf. Acesso em: 03/11/2021).

⁵⁵² LESSIG, Lawrence. **Code**. Version 2.0. New York: Basic Books, 2006. p. 125-137, Disponível em: https://cyber.harvard.edu/ptc2010/sites/ptc2010/images/Lessig_Code_Excerpts.pdf. Acesso em: 04/12/2021.

⁵⁵³ Para uma melhor compreensão sobre os sistemas de moderação de conteúdo em escala, conferir: SUZOR, Nicolas. Understanding content moderation systems: new methods to understand internet governance at scale, over time, and across platforms. In: WHALEN, Ryan (Ed.) **Computational Legal Studies: The Promise and Challenge of Data-Driven Research**. Edward Elgar Publishing, p. 166-189, 2020. Disponível em: <https://eprints.qut.edu.au/129464/16/Suzor%2B2018%2BUnderstanding%2Bcontent%2Bmoderation%2Bsystem%2Bclean.pdf>. ou <https://doi.org/10.4337/9781788977456.00013>. Acesso em: 20/07/2022. Ver também: MUSIANI, Francesca. Governance algoritmica: sorveglianza, censura e diritti fondamentali. In: FOSSA, Fabio; SCHIAF-FONATI, Viola; TAMBURRINI, Guglielmo (Eds.). **Automi e persone. Introduzione all’etica dell’intelligenza artificiale e della robotica**. Carocci, p. 95-113, 2021. Disponível em: <https://halshs.archives-ouvertes.fr/halshs-03381546>. Acesso em: 23/07/2022.

⁵⁵⁴ Tradução livre do original: “[...] automated moderation threatens to (a) further increase opacity, making a famously non-transparent set of practices even more difficult to understand or audit, (b) further complicate outstanding issues of fairness and justice in large-scale sociotechnical systems and (c) re-obscure the fundamentally political nature of speech decisions being executed at scale.” (GORWA, Robert; BINNS, Reuben; KATZENBACH, Chistian. Algorithmic content moderation: Technical and political challenges in the

(iv) *A capacidade da comunidade de autocorreção*: nesse aspecto, leva em consideração a comunidades fechadas a facilidade de uma autocorreção entre os próprios usuários, de forma que os remédios apresentados pode auxiliá-los nesse propósito. A exemplo, a *Wikipédia* possui um autopolicimento, proporcionando uma abertura e aumento de possibilidades de autocorreção.⁵⁵⁵

(v) *Como os remédios afetam outras pessoas*: importante os moderadores, ao aplicarem os remédios, terem conhecimento dos seus efeitos. Sobretudo, as necessidades de vítimas específicas, e dos usuários em geral, com o objetivo de aplicar os remédios e compatibilizar seus interesses de acordo com o modelo de negócio, as consequências às vítimas e o público em geral, procurando compreender as necessidades das partes afetadas. Como bem explica Goldman, “em algumas circunstâncias, a remoção é a única opção sustentável para eliminar o dano. Em outros, os remédios de não remoção ajudam a equilibrar os interesses concorrentes/conflitantes, como a liberdade de expressão do autor, enquanto ainda beneficiam partes externas.”⁵⁵⁶

(vi) *Rever o envolvimento do usuário, enquanto restringe as violações e reincidências*: as plataformas têm o dilema de coibir violações a suas regras, e, por outro lado, não perder um usuário (consumidor) de seus produtos. Com isso, devem procurar remédios que atendam ambos os propósitos: não banir ou remover o usuário/conta, e, ao mesmo tempo, evitar a reincidência nas violações, contudo, permanecendo na plataforma.

(vii) *Sanções paralelas*: esse fator influencia na tomada de decisão, quanto à avaliação do moderador em adotar um remédio que possa surtir efeitos fora das plataformas, causando outras sanções ao usuário – judicial ou extrajudicialmente. Dessa forma, deve ser examinado o melhor remédio, para não haver desproporcionalidade entre a violação na plataforma, e as sanções fora do ambiente *online*, por exemplo, rescisão de contrato de trabalho, processo judicial, afastamento do emprego etc.

Assim, finalizamos a demonstração dos remédios alternativos, não sendo o objetivo exaurir todos os recursos existentes, mas realizar um panorama dos principais, para que sejam

automation of platform governance. **Big Data & Society**. 2020. p. 1. Disponível em: <https://doi.org/10.1177/2053951719897945>. Acesso em: 20/07/2022).

⁵⁵⁵ GOLDMAN, Eric. Content Moderation Remedies. **Michigan Technology Law Review**. v. 28, n. 1. p. 1-59, 2021. p. 45. Disponível em: <https://repository.law.umich.edu/mtlr/vol28/iss1/2>. Acesso em: 02/11/2021.

⁵⁵⁶ Tradução livre do original: “In some circumstances, removal is the only tenable option to eliminate the harm. In others, non-removal remedies help balance the competing/conflicting interests, such as the author’s free expression, while still benefitting external parties.” (GOLDMAN, Eric. Content Moderation Remedies. **Michigan Technology Law Review**. v. 28, n. 1. p. 1-59, 2021. p. 46. Disponível em: <https://repository.law.umich.edu/mtlr/vol28/iss1/2>. Acesso em: 02/11/2021).

ferramentas utilizadas pelos moderadores de conteúdo, com o afã de remodelar a ideia de uma moderação binária, indo além da remoção, traçando-se alguns fatores importantes a serem considerados quanto a aplicação dos remédios.

É cediço que o caminho da moderação de conteúdo, ainda é muito longo, apenas começamos. O debate da regulação do discurso *online*, é relativamente novo, apesar da grande construção doutrinária. Entretanto, as práticas, os erros, as soluções, as diretrizes governamentais e da organização civil, que tem construído esse universo da moderação de conteúdo, ainda carecem de muito estudo e aperfeiçoamento, sobretudo quanto a transparência.

Nesse sentido, como forma de apontar o futuro da moderação de conteúdo, abordamos dois pontos importantes, os quais repaginam e colocam novas perspectivas sobre a moderação de conteúdo pelas plataformas. Não se pretende, realizar um estudo completo sobre o tema, mas, tão somente, traçar nossas perspectivas e pontos de inflexão para estudos futuros.

a) *Recalibragem da moderação de conteúdo*: esse primeiro ponto, diz respeito às condutas que as plataformas *online* tiveram de adotar, em decorrência da pandemia causada pelo COVID-19,⁵⁵⁷ demonstrando uma nova realidade da moderação de conteúdo⁵⁵⁸. Com as recomendações da OMS sobre a necessidade de restrição das atividades, objetivando evitar a propagação do vírus, as principais plataformas dispensaram seus funcionários, passando a realizar a moderação de conteúdo por meio de ferramentas de inteligência artificial – IA.⁵⁵⁹

⁵⁵⁷ Importante deixar assente, que até a conclusão deste trabalho, não foi declarada o fim da pandemia causada pelo COVID-19, pela Organização Mundial de Saúde – OMS.

⁵⁵⁸ Importante o apontamento de Evelyn Douek: “A moderação de conteúdo durante o COVID-19 representou, portanto, uma mudança retórica dramática das plataformas, mas não uma mudança verdadeiramente substantiva. De certa forma, simplesmente representou um momento incomum de franqueza sobre a lógica governante do design do sistema de moderação de conteúdo.” Tradução livre do original: “Content moderation during COVID-19 therefore represented a dramatic rhetorical shift from platforms, but not a truly substantive one. In some ways, it simply represented an unusual moment of candor about the governing logics of content moderation system design.” (DOUEK, Evelyn. *Governing online speech: from “posts-as-trumps” to proportionality and probability*. Columbia Law Review. v. 121, n. 3, p. 759-837, 2021. p. 803. Disponível em: https://www.columbialawreview.org/wp-content/uploads/2021/04/Douek-Governing_Online_Speech-from_Posts_As-Trusts_To_Proportionality_And_Probability.pdf. Acesso em: 03/11/2021).

⁵⁵⁹ O Facebook declarou: “[Com] uma força de trabalho reduzida e remota, agora dependeremos mais de nossos sistemas automatizados para detectar e remover conteúdo violador e desabilitar contas. Como resultado, esperamos cometer mais erros e as revisões levarão mais tempo do que o normal, mas continuaremos monitorando o desempenho de nossos sistemas e fazendo ajustes.” Da mesma forma, o Twitter noticiou que estaria “aumentando nosso uso de aprendizado de máquina e automação para realizar uma ampla gama de ações em conteúdo potencialmente abusivo e manipulador. Queremos ser claros: enquanto trabalhamos para garantir que nossos sistemas sejam consistentes, às vezes eles podem não ter o contexto que nossas equipes trazem, e isso pode resultar em erros.” Na mesma linha, o YouTube pronunciou: “Nós começaremos temporariamente a confiar mais na tecnologia para ajudar em parte do trabalho normalmente feito pelos revisores. Isso significa que os sistemas automatizados começarão a remover alguns conteúdos sem revisão humana, para que possamos continuar a agir rapidamente para remover conteúdo violador e proteger nosso ecossistema, enquanto temos

Desde então, percebeu-se uma abordagem bastante agressiva nessa moderação, diferentemente da abordagem até então adotada.⁵⁶⁰ Com a severa remoção de notícias falsas, conteúdos de exploração, passaram a veicular informações confiáveis, conforme orientações da Organização Mundial da Saúde.

Verificou-se, que a moderação mais rígida, sobretudo quando ao combate à desinformação – no contexto pandêmico o dano potencial de informações falsas tem uma conotação mais direta, podendo levar a morte – ficando conhecida como a “Infodemia COVID-19”,⁵⁶¹ ao lado da moderação algorítmica, permitiu uma nova postura, inclusive, com as plataformas reconhecendo o fato dessas novas regras, aplicadas pela inteligência artificial, levarem ao aumento de “falsos positivos” – como descrito acima, significa a remoção de conteúdo que deveria permanecer publicado. Com isso, notou-se o aumento da utilização do padrão de “proporcionalidade” na moderação do conteúdo, por parte das plataformas, permitindo um exame sob a violação do discurso, que seja equilibrada com seus benefícios, chamado por Zittrain de “ascensão da Era da Saúde Pública”⁵⁶² da governança da *Internet*, que analisa o custo-benefício de diferentes regras, e não apenas quanto aos direitos dos usuários – orador.

Um outro comportamento perceptível das plataformas, foi o inédito reconhecimento da possibilidade de “erros” surgidos inevitavelmente do uso das ferramentas de IA, pois, as empresas sempre mantiveram uma postura de defender o uso da automação em escala, como um recurso infalível na moderação do conteúdo.

Nesse contexto, a moderação de conteúdo entrou num “estado de emergência” durante a pandemia de COVID-19,⁵⁶³ demonstrando uma nova roupagem ou modo de visão da moderação, proporcionando uma recalibração no entendimento até então vigente, como explica Douek, por entender que devemos agora pensar não mais numa moderação de conteúdo por meio de um olhar individualista específico, devendo as plataformas, os

proteções no local de trabalho. À medida que fazemos isso, usuários e criadores podem ver um aumento nas remoções de vídeos.” (DOUEK, Evelyn, COVID-19 and Social Media Content Moderation, **LAWFARE**, 2020. Disponível em: <https://www.lawfareblog.com/covid-19-and-social-media-content-moderation>. Acesso em: 20/07/2022)

⁵⁶⁰ DOUEK, Evelyn, COVID-19 and Social Media Content Moderation, **LAWFARE**, 2020. Disponível em: <https://www.lawfareblog.com/covid-19-and-social-media-content-moderation>. Acesso em: 20/07/2022.

⁵⁶¹ YANG, Kai-Cheng; PIERRI, Francesco; HUI, Pik-Mai [et al.]. “The COVID-19 Infodemic: Twitter versus Facebook. **Big Data & Society**, 2021. Disponível em: <https://doi.org/10.1177/20539517211013861>. Acesso em: 21/07/2022.

⁵⁶² ZITTRAIN, Jonathan. **Three Eras of Digital Governance**. Disponível em: <https://ssrn.com/abstract=3458435> ou <http://dx.doi.org/10.2139/ssrn.3458435>. Acesso em: 08/07/2022. Abordamos o tema no item 3.1.

⁵⁶³ DOUEK, Evelyn, COVID-19 and Social Media Content Moderation, **LAWFARE**, 2020. Disponível em: <https://www.lawfareblog.com/covid-19-and-social-media-content-moderation>. Acesso em: 20/07/2022.

reguladores e os usuários em geral reconhecerem essa abordagem tendo teve seu auge no início da *internet* e não mais se sustentando. Com essa nova realidade, as plataformas têm se mostrado empenhadas em equilibrar os interesses sociais e tendentes a escolher entre os custos do erro de forma sistêmica.⁵⁶⁴

Essa recalibragem é urgente por estarmos vivenciando um novo estágio da governança do discurso *online*, com as plataformas à frente desse desenvolvimento, criando meios de atender aos anseios da sociedade por uma maior responsabilidade na moderação de conteúdo, seja por meio de inovações ou reformas autorregulatórias, a exemplo do supracitado *Facebook Oversight Board*, demonstrando uma recalibragem rumo a um olhar mais sistêmico, evoluindo de uma concepção categórica e individualista de direitos de fala *online*, para uma concepção em torno de dois preceitos diferentes: *proporcionalidade* e *probabilidade*. Como defendido por Douek, sob esses preceitos a moderação de conteúdo é uma questão de “*equilíbrio sistêmico*: as regras são escritas para abranger múltiplos interesses, não apenas direitos de fala individuais, e com consciência das taxas de erro inerentes à aplicação de qualquer regra na escala verdadeiramente impressionante das principais plataformas.”⁵⁶⁵

Com essa nova perspectiva, nota-se a distinção nas decisões pautadas na proporcionalidade e probabilidade, sendo, como explica Douek, a *proporcionalidade* exigindo a justificação da intervenção nos direitos e, quanto mais contundente for essa intervenção, maiores e mais fortes serão as justificativas exigidas. Com isso, endereça aos moderadores um equilíbrio entre os interesses sociais e os interesses individuais. Enquanto a concepção da *probabilidade* do discurso *online*, já reconhece não ser perfeita a aplicação dessas regras elaboradas como resultado desse balanceamento e, com isso, os sistemas de governanças devam ter a percepção de ser inevitável a ocorrência de erros e que devem escolher os níveis e tipos de erros aceitáveis. Em resumo, Douek afirma que a aceitação consciente do fato de que as “determinações de fala erradas em algumas porcentagens de casos são inerentes à

⁵⁶⁴ DOUEK, Evelyn. Governing online speech: from “posts-as-trumps” to proportionality and probability. *Columbia Law Review*. v. 121, n. 3, p. 759-837, 2021. p. 762. Disponível em: https://www.columbialawreview.org/wp-content/uploads/2021/04/Douek-Governing_Online_Speech-from_Posts_As-Trumps_To_Proportionality_And_Probability.pdf. Acesso em: 03/11/2021

⁵⁶⁵ Tradução livre do original: “[...] systemic balancing: Rules are written to encompass multiple interests, not just individual speech rights, and with awareness of the error rates inherent in enforcing any rule at the truly staggering scale of major platforms.” (DOUEK, Evelyn. Governing online speech: from “posts-as-trumps” to proportionality and probability. *Columbia Law Review*. v. 121, n. 3, p. 759-837, 2021. p. 763. Disponível em: https://www.columbialawreview.org/wp-content/uploads/2021/04/Douek-Governing_Online_Speech-from_Posts_As-Trumps_To_Proportionality_And_Probability.pdf. Acesso em: 03/11/2021)

governança de fala *online* requer ser muito mais sincero sobre as taxas de erro, o que pode permitir a calibração da criação de regras para as realidades práticas de aplicação.”⁵⁶⁶

Com efeito, a pandemia do COVID-19 descortinou a necessidade de um equilíbrio sistêmico na governança do discurso *online*, deixando de lado, uma atuação mais individualizada da moderação de conteúdo. Entendemos, que nesse novo pensar sistêmico, não se deve desvalorizar ou enfraquecer o direito da liberdade de expressão, devido essa abordagem tender a não exigência de resultados substantivos particulares, sendo importante, como lembra Douek, reconhecer que esse papel do equilíbrio sistêmico orienta a discurso *online* em torno das questões certas.⁵⁶⁷ Portanto, apontamos essa recalibragem ter direcionado meios de compreender como deve ser a moderação de conteúdo em um mundo pós-pandemia.

b. *O futuro da moderação de conteúdo: do formalismo para o realismo.* Esse segundo ponto deriva da constatação do primeiro, da necessidade de reconhecer uma moderação de conteúdo não mais individualizada e específica, adotando um equilíbrio sistêmico. Com isso, a moderação de conteúdo praticada hoje fora moldada sob o aspecto formalista, o qual, diante da vastidão de variáveis a serem moderadas, tornou-se ineficaz⁵⁶⁸. As plataformas decidem entre agir ou não em milhões de conteúdos carregados todos os dias, derivando num complexo sistema burocrático de decisões sobre como os indivíduos pode usar o placó da *internet*. Douek afirma que o “modelo formalista, invocando normas de raciocínio e precedentes de estilo judicial, está fadado ao fracasso nessa escala e complexidade.”⁵⁶⁹

⁵⁶⁶ Tradução livre do original: “[...] speech determinations wrong in some percentage of cases is inherent in online speech governance requires being much more candid about error rates, which can allow for the calibration of rulemaking to the practical realities of enforcement.” (DOUEK, Evelyn. Governing online speech: from “posts-as-trumps” to proportionality and probability. **Columbia Law Review**. v. 121, n. 3, p. 759-837, 2021. p. 764. Disponível em: https://www.columbialawreview.org/wp-content/uploads/2021/04/Douek-Governing_Online_Speech-from_Posts_As-Trumps_To_Proportionality_And_Probability.pdf. Acesso em: 03/11/2021).

⁵⁶⁷ DOUEK, Evelyn. Governing online speech: from “posts-as-trumps” to proportionality and probability. **Columbia Law Review**. v. 121, n. 3, p. 759-837, 2021. p. 769. Disponível em: https://www.columbialawreview.org/wp-content/uploads/2021/04/Douek-Governing_Online_Speech-from_Posts_As-Trumps_To_Proportionality_And_Probability.pdf. Acesso em: 03/11/2021.

⁵⁶⁸ Eficácia no sentido de que a moderação de conteúdo realizada de forma eficaz, é capaz de construir uma situação de fala online ideal. Nesse sentido, Evelyn Douek aduz que a eficácia “significa a aplicação eficiente das regras da plataforma de uma forma que aumenta a responsabilidade pública e a transparência, o que cria um mecanismo para que esses stakeholders influenciem as operações das plataformas ou resolvam os problemas criados por eles.” Tradução livre do original: “[...] it means the efficient enforcement of platform rules in a way that increases public accountability and transparency, which creates a mechanism for these stakeholders to influence platforms’ operations or address the problems created by them.” (DOUEK, Evelyn, *The Siren Call of Content Moderation Formalism*. In: BOLLIGER, Lee; GEOFFREY, Stone (Eds.). **New Technologies of Communication and the First Amendment: The Internet, Social Media and Censorship**. Oxford Press, forthcoming 2022. p. 4. Disponível em: <https://ssrn.com/abstract=4005314> ou <http://dx.doi.org/10.2139/ssrn.4005314>. Acesso em: 20/07/2022).

⁵⁶⁹ Tradução livre do original: “A formalistic model, invoking judicial-style norms of reasoning and precedent, is doomed to fail at this scale and complexity.” (DOUEK, Evelyn, *The Siren Call of Content Moderation Formalism*. In: BOLLIGER, Lee; GEOFFREY, Stone (Eds.). **New Technologies of Communication and the**

O formalismo, sem dúvidas, tem suas virtudes, fixando a ideia de previsibilidade e consistência típicos do Estado de Direito, servindo para orientar o comportamento do usuário a não violar as regras. Além disso, a fixação de regras padronizadas e o compromisso das plataformas em aplicá-las de forma semelhante aos demais casos, restringe a discricionariedade dos moderadores. Contudo, tal sistema tem se mostrado incapaz de alcançar a moderação moderna. Pois, nessa construção formalista, exige-se cada vez mais detalhes peculiares e sutis de comportamentos, levando a acomodar e categorizar uma infinita variedade de conteúdo *online*, impossível de ser prevista e descrita nas regras internas.⁵⁷⁰

Nesse sentido, Douek leciona que atualmente as práticas da moderação de conteúdo tem indicado o formalismo estando ainda mais fora de alcance neste contexto, em relação a outros e acrescenta: “o mundo *offline* ao qual a governança de moderação de conteúdo é frequentemente comparada pudesse alcançá-lo, a escala e a confusão da governança de fala *online* tornam a tradução simples de ferramentas *offline* inadequada.”⁵⁷¹

Nessa perspectiva, constata-se, conforme estudo realizado por Douek, que devemos trilhar do formalismo para o realismo de moderação de conteúdo, buscando um sistema melhor e mais legítimo de governança de moderação de conteúdo, que “não deve ser abandonado apenas porque o modelo formalista que as partes interessadas aspiram atualmente é impossível, mas ser realista sobre o que almejar é uma parte necessária da construção desse sistema.”⁵⁷² Constitui, como parte integrante desse sistema, a necessidade de desenvolver uma compreensão de aceitar o que pode e o que não pode ser restringido, de ter a noção de quais erros podem e não podem ser remediados e as compensações inerentes a cada escolha. Para a

First Amendment: The Internet, Social Media and Censorship. Oxford Press, forthcoming 2022. p. 2. Disponível em: <https://ssrn.com/abstract=4005314> ou <http://dx.doi.org/10.2139/ssrn.4005314>. Acesso em: 20/07/2022).

⁵⁷⁰ Nessa ideia, a “moderação de conteúdo precisa de uma forma de governança mais fluida, experimental e responsiva do que o formalismo oferece.” Tradução livre do original: “Content moderation needs a more fluid, experimental and responsive form of governance than formalism provides.” (DOUEK, Evelyn, The Siren Call of Content Moderation Formalism. In: BOLLIGER, Lee; GEOFFREY, Stone (Eds.). **New Technologies of Communication and the First Amendment:** The Internet, Social Media and Censorship. Oxford Press, forthcoming 2022. p. 16. Disponível em: <https://ssrn.com/abstract=4005314> ou <http://dx.doi.org/10.2139/ssrn.4005314>. Acesso em: 20/07/2022).

⁵⁷¹ Tradução livre do original: “[...] offline world to which content moderation governance is frequently compared could achieve it, the scale and messiness of online speech governance makes simple translation of offline tools inapposite.” (DOUEK, Evelyn, The Siren Call of Content Moderation Formalism. In: BOLLIGER, Lee; GEOFFREY, Stone (Eds.). **New Technologies of Communication and the First Amendment:** The Internet, Social Media and Censorship. Oxford Press, forthcoming 2022. p. 4. Disponível em: <https://ssrn.com/abstract=4005314> ou <http://dx.doi.org/10.2139/ssrn.4005314>. Acesso em: 20/07/2022).

⁵⁷² Tradução livre do original: “[...] should not be forsaken just because the formalistic model that stakeholders currently aspire to is impossible, but being realistic about what to aim for is a necessary part of building that system.” (DOUEK, Evelyn, The Siren Call of Content Moderation Formalism. In: BOLLIGER, Lee; GEOFFREY, Stone (Eds.). **New Technologies of Communication and the First Amendment:** The Internet, Social Media and Censorship. Oxford Press, forthcoming 2022. p. 4. Disponível em: <https://ssrn.com/abstract=4005314> ou <http://dx.doi.org/10.2139/ssrn.4005314>. Acesso em: 20/07/2022).

constituição de um modelo de moderação de conteúdo com maior eficácia, é preciso sacrificar algum grau de consistência e determinação.

A obtenção desse novo sistema requer um grau de experimentalismo institucional. Contudo, Douek expõe que alguns princípios gerais devem orientar o experimentalismo institucional da moderação de conteúdo, para não permitirmos a operação dessas regras no vácuo, devendo obediência a um contexto institucional, político e social, sem erradicar a discricionariedade e aceitando o fato de sempre haver erros a serem considerados em sua aplicação. A pesquisadora formula os seguintes princípios: 1. *Capacitar os tomadores de decisão que não sejam advogados*; 2. *Abandone a aspiração a regras universalmente consistentes*; 3. *Esperar mudanças e adaptações constantes*; 4. *Fale sobre os erros de forma diferente*; 5. *Planeje e canalize o desacordo*.⁵⁷³

Portanto, reputamos importante o segundo ponto na construção de uma moderação de conteúdo próxima do ideal, inicialmente transcendendo de uma visão individualizada para um equilíbrio sistêmico e, finalmente, migrando de um modelo formalista, para um realista, considerando a desordem que é a fala *online* em plataformas, num mundo vasto e em permanente alteração e evolução, dificultando uma moderação consistente em regras rígidas.

3.4. Aspectos normativos da moderação de conteúdo no Brasil

A governança da *internet* no Brasil foi reconhecida internacionalmente por adotar um modelo pluriparticipativo centrado em um Comitê Gestor da *Internet* (CGI)⁵⁷⁴ e, conforme dispõe o Decreto Presidencial n. 4.829/2003, é o responsável pelo estabelecimento de diretrizes estratégicas relacionadas ao funcionamento, ao desenvolvimento e ao uso da

⁵⁷³ Tradução livre do original: “Principle 1 – Empower decisionmakers other than lawyers; Principle 2 – Abandon the aspiration to universally consistent rules; Principle 3 – Expect constant change and adaptation; Principle 4 – Talk about errors differently; Principle 5 – Plan for and channel disagreement.” (DOUEK, Evelyn, *The Siren Call of Content Moderation Formalism*. In: BOLLIGER, Lee; GEOFFREY, Stone (Eds.). **New Technologies of Communication and the First Amendment: The Internet, Social Media and Censorship**. Oxford Press, forthcoming 2022. p. 17-22. Disponível em: <https://ssrn.com/abstract=4005314> ou <http://dx.doi.org/10.2139/ssrn.4005314>. Acesso em: 20/07/2022).

⁵⁷⁴ O CGI foi criado em 1995, a atual composição do CGI.br foi estabelecida pelo Decreto Nº 4.829, de 3 de setembro de 2003, da Presidência da República. Integram o CGI.br: 9 representantes do setor governamental, 4 do setor empresarial, 4 do terceiro setor, 3 da comunidade científica e tecnológica e 1 representante de notório saber em assuntos de Internet. (COMITÊ GESTOR DA INTERNET NO BRASIL. Composição do CGI.br. **CGI.br**. 2022. Disponível em: <https://www.cgi.br/membros/>. Acesso em: 01/08/2022).

Internet no Brasil, por meio de resoluções resultantes do processo deliberativo que busca o consenso dos conselheiros.⁵⁷⁵

Nesse ponto, fruto importante dessa construção colaborativa do CGI, o Decálogo, adotado em 2009, engloba os seguintes enfoques:

1. Liberdade, privacidade e direitos humanos: O uso da Internet deve guiar-se pelos princípios de liberdade de expressão, de privacidade do indivíduo e de respeito aos direitos humanos, reconhecendo-os como fundamentais para a preservação de uma sociedade justa e democrática. **2. Governança democrática e colaborativa:** A governança da Internet deve ser exercida de forma transparente, multilateral e democrática, com a participação dos vários setores da sociedade, preservando e estimulando o seu caráter de criação coletiva. **3. Universalidade:** O acesso à Internet deve ser universal para que ela seja um meio para o desenvolvimento social e humano, contribuindo para a construção de uma sociedade inclusiva e não discriminatória em benefício de todos. **4. Diversidade:** A diversidade cultural deve ser respeitada e preservada e sua expressão deve ser estimulada, sem a imposição de crenças, costumes ou valores. **5. Inovação:** A governança da Internet deve promover a contínua evolução e ampla difusão de novas tecnologias e modelos de uso e acesso. **6. Neutralidade da rede:** Filtragem ou privilégios de tráfego devem respeitar apenas critérios técnicos e éticos, não sendo admissíveis motivos políticos, comerciais, religiosos, culturais, ou qualquer outra forma de discriminação ou favorecimento. **7. Inimputabilidade da rede:** O combate a ilícitos na rede deve atingir os responsáveis finais e não os meios de acesso e transporte, sempre preservando os princípios maiores de defesa da liberdade, da privacidade e do respeito aos direitos humanos. **8. Funcionalidade, segurança e estabilidade:** A estabilidade, a segurança e a funcionalidade globais da rede devem ser preservadas de forma ativa através de medidas técnicas compatíveis com os padrões internacionais e estímulo ao uso das boas práticas. **9. Padronização e interoperabilidade:** A Internet deve basear-se em padrões abertos que permitam a interoperabilidade e a participação de todos em seu desenvolvimento. **10. Ambiente legal e regulatório:** O ambiente legal e regulatório deve preservar a dinâmica da Internet como espaço de colaboração.⁵⁷⁶

Nesse contexto, a aceitação e o teor do Decálogo serviram de propulsores para a elaboração do Marco Civil para a *Internet* no Brasil.⁵⁷⁷ É bem verdade, o projeto de lei, que originou a Lei n. 12.965/2014, decorrer de uma resposta de várias *stakeholders*, a tentativas

⁵⁷⁵ BRASIL. Decreto nº 4.829, de 3 de setembro de 2003. Dispõe sobre a criação do Comitê Gestor da Internet no Brasil - CGI.br, sobre o modelo de governança da Internet no Brasil, e dá outras providências. **Diário Oficial da República Federativa do Brasil**, Poder Executivo. Brasília, DF, 4 de setembro de 2003.

⁵⁷⁶ COMITÊ GESTOR DA INTERNET NO BRASIL. Princípios para o Uso e a Governança da Internet no Brasil. Resolução CGI.br/RES/2009/003/P. **CGI.br**. 2009. Disponível em: www.cgi.br/resolucoes/documento/2009/003. Acesso em: 30/07/2022.

⁵⁷⁷ Quanto ao contexto de aprovação do Marco Civil da Internet, Paulo Gonet Branco e Pedro Henrique Gonet Branco, sublinham que “o projeto de lei foi aprovado em um contexto peculiar. Chegou ao conhecimento de autoridades brasileiras e a opinião pública a alegação de Edward Snowden de que a NSA programa de vigilância estava monitorando o sistema de telecomunicações do Brasil e que a política e informações econômicas sensíveis estavam sendo coletadas. Este foi um forte impulso para o Brasil promulgar o Marco Civil da Internet (MCI) de 2014.” Tradução livre do original: “The bill passed in a peculiar context. It came to the attention of Brazilian authorities and public opinion the Edward Snowden allegation that NSA’s surveillance program was monitoring Brazil’s telecommunication system and that political and economic sensitive information were being gathered. This was a strong push to Brazil enact the 2014 Marco Civil da Internet (MCI).” (BRANCO, Paulo Gustavo Gonet; BRANCO, Pedro Henrique de Moura Gonet. *Structural Way of Governing the Internet – The Brasilia Position*. **Revista Jurídica**. v. 2, n. 55, p. 592-606, Curitiba, 2019. p. 600. Disponível em: <http://revista.unicuritiba.edu.br/index.php/RevJur/article/view/3840/371372180>. Acesso em: 06/11/2021).

de criminalizar condutas pertinentes à *Internet*, sem que antes houvesse um rol de direitos e garantias fundamentais da *Internet*, no ordenamento jurídico brasileiro.⁵⁷⁸ A partir disso, durante os anos de 2009 a 2011, houve um processo colaborativo, aberto, com ampla participação de diversos atores, por meio de consultas públicas *online* e presenciais, culminando com a remessa, no final de 2011, de um projeto de lei à Câmara dos Deputados. Após a tramitação, a lei do Marco Civil da Internet (Lei n. 12.965/2014) foi aprovada no Congresso e promulgada pela então presidente Dilma Rousseff, na abertura do Encontro Multissetorial Global sobre o Fundo da Governança da *Internet* – NetMundial, realizado na cidade de São Paulo, em abril de 2014.⁵⁷⁹

3.4.1. O Marco Civil da Internet

A moderação de conteúdo no Brasil pelas grandes plataformas de mídia social, assume aspectos distintos e próprios, sobretudo em decorrência da previsão normativa regida pelo Marco Civil da Internet (Lei n. 12.965/2014). Conforme a própria descrição, a Lei afirma que “estabelece princípios, garantias, direitos e deveres para o uso da *Internet* no Brasil”. Dispõe, ao longo dos Artigos 18 a 21, regras atinentes à governança da *Internet*, estipulando incentivos para evitar a derrubada de conteúdo do ambiente *online*. Importante frisar que aludida norma não dispõe, exclusivamente, sobre a moderação de conteúdos por plataformas de redes sociais, englobando, também, outros intermediários, como sites de buscas, blog ou sites de notícias, plataformas de compras com interação de usuários em comentários etc.

Dentre esses comandos normativos, o Artigo 19⁵⁸⁰ merece especial atenção, pois dispõe no *caput* a prioridade de assegurar a liberdade de expressão e impedir a censura, estabelecendo sobre a responsabilidade civil dos provedores de aplicações, em decorrência de conteúdos criados por terceiros, somente incidir após ordem judicial específica para a sua.

⁵⁷⁸ WAGNER, Flávio Rech; CANABARRO, Diego Rafael. A governança da Internet: definição, desafios e perspectivas. In: PIMENTA, Marcelo Soares; CANABARRO, Diego Rafael (Org.). **Governança Digital**. Porto Alegre: Editora da UFRGS/CEGOV, 2014. p. 203.

⁵⁷⁹ Para um melhor estudo sobre o processo de elaboração colaborativo do Marco Civil da Internet no Brasil, consultar: CRUZ, Francisco Carvalho de Brito. **Direito, democracia e cultura digital**: a experiência de elaboração legislativa do Marco Civil da Internet. 2015. Dissertação (Mestrado em Filosofia e Teoria Geral do Direito) - Faculdade de Direito, Universidade de São Paulo, São Paulo, 2015; e SANTARÉM, Paulo Rená da Silva. **O direito achado na rede**: a emergência do acesso à Internet como direito fundamental no Brasil. 2010. 158 f., il. Dissertação (Mestrado em Direito, Estado e Constituição) - Universidade de Brasília, Brasília, 2010.

⁵⁸⁰ Confira o inteiro teor: “Art. 19. 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. [...]”

Desta forma, assume um regime de responsabilidade civil favorável a intermediários, separando o autor da postagem de seu mensageiro. Extrai-se que o dispositivo parte do pressuposto de as plataformas não atuarem como moderadoras dessa temática, gerenciando, moldando e monitorando toda a produção desse conteúdo gerado em seus domínios.

Nessa linha, o modelo brasileiro encampa, de certa forma, os *Princípio de Manila* para a responsabilidade de intermediários,⁵⁸¹ conforme se percebe da redação do item I “Os intermediários devem ser protegidos da responsabilidade por conteúdo de terceiros”, e apresenta quatro recomendações (a) “quaisquer regras que regulem a responsabilidade do intermediário devem estar previstas em lei, que deve ser preciso, claro e acessível;” Bem como, (b) “os intermediários devem ser isentos de responsabilidade por conteúdo de terceiros em circunstâncias em que não tenham estado envolvidos na modificação desse conteúdo;” (c) os intermediários não devem ser responsabilizados por não restringir conteúdo lícito;” Além disso, (d) “os intermediários nunca devem ser responsabilizados estritamente por hospedar conteúdo ilegal de terceiros, nem devem ser obrigados a monitorar o conteúdo de forma proativa como parte de um regime de responsabilidade do intermediário.”⁵⁸²

Acrescente-se à análise o referido Artigo 19 da Marco Civil da Internet, como objeto de dois Recursos Extraordinários no Supremo Tribunal Federal, no qual será decidida a constitucionalidade desse modelo de responsabilidade civil. Até a conclusão deste trabalho, os RE n. 1.057.25/MG e 1.037.396/SP não foram pautados para julgamento. O primeiro caso faz menção a fatos anteriores ao Marco Civil da Internet, o segundo, por sua vez, indaga se o Art. 19 do Marco Civil da Internet fere as garantias constitucionais, como direitos a personalidade, acesso à justiça, dentre outros.⁵⁸³

⁵⁸¹ Ver item 3.2, desse Capítulo.

⁵⁸² Tradução livre do original “I. Intermediaries should be shielded from liability for third-party content. a) Any rules governing intermediary liability must be provided by laws, which must be precise, clear, and accessible. b) Intermediaries should be immune from liability for third-party content in circumstances where they have not been involved in modifying that content. c) Intermediaries must not be held liable for failing to restrict lawful content. d) Intermediaries must never be made strictly liable for hosting unlawful third-party content, nor should they ever be required to monitor content proactively as part of an intermediary liability regime.” (**THE SANTA CLARA PRINCIPLES: On Transparency and Accountability in Content Moderation**. Disponível em: <https://santaclaraprinciples.org/>. Acesso em: 12/07/2022))

⁵⁸³ Laux, resume que “no âmbito do Poder Judiciário, ganha enorme relevo atual a discussão inserta pelo RE n. 1.037.396/SP, em trâmite no Supremo Tribunal Federal, com repercussão geral admitida e que discute justamente a constitucionalidade do art. 19 do Marco Civil da Internet. O pedido recursal apresentado é para que seja reconhecida a constitucionalidade do dispositivo que impõe responsabilidade civil aos provedores de hospedagem de postagens de terceiros somente na hipótese prevista em lei, ou seja, após a configuração de conduta omissa em relação a uma determinação judicial de remoção de conteúdo. A “via alternativa”, sugerida pela parte recorrida, seria a do restabelecimento do sistema prévio e altamente casuístico, vinculado a uma necessidade de apreciação, pelo particular, da licitude de expressões postadas na internet, inclusive para fins de responsabilização em caso de manutenção do material na rede.” (LAUX, Francisco de Mesquita. Modelos de responsabilização civil por conteúdo postado na internet: Direito Comparado, Marco Civil e discussões nos

Ademais, o regime de responsabilidade civil possui algumas exceções, desta forma, o que dispõe o Art. 21 do Marco Civil da Internet, afirma que a responsabilidade civil da plataforma será subsidiária pela não remoção, depois de simples notificação, em casos de violação da intimidade por materiais contendo imagens não autorizadas de nudez, atos sexuais privados e demais conteúdo dessa natureza. Uma outra exceção, são os casos que envolvem de direitos autorais, previstos no Art. 19, § 2º do Marco Civil da Internet, o qual afirma a aplicação desta Lei somente nas situações que ocorreram após a promulgação de legislação específica, porém, ainda não aconteceu. Assim, nessas circunstâncias, continua em vigor o regime anterior ao Marco Civil da Internet, ou seja, o sistema de responsabilização civil a partir da notificação. Contudo, entendemos as plataformas com o poder de realizar essa moderação, mesmo sem a edição dessa lei específica, por meio de sua governança privada.

Analizando, até aqui, o Marco Civil da Internet, verifica-se sua relação lógica direta à governança estatal, silenciando completamente quanto à governança privada pelos intermediários. Sendo assim, o supracitado Marco Civil da Internet possui a proposição da prerrogativa exclusiva do poder judiciário para a remoção de conteúdos. Esse fato deu-se porque, na época da aprovação do Marco Civil, não havia enfoque nos conteúdos restringidos pela moderação de conteúdo, por parte dos provedores de aplicação de *internet*.⁵⁸⁴ Pretendia-se, desse modo, evitar, devido à atuação do judiciário, que os provedores de *internet* opusessem-se contra os conteúdos gerados por seus usuários para evitar de serem responsabilizados por danos causados por eles.⁵⁸⁵

Nesse ponto, em especial, não encontramos problemas quanto à função “tradicional” do Poder Judiciário em dirimir conflitos relacionados à legalidade ou a ilegalidade de um determinado conteúdo, exercendo sua lógica binária: permitir ou proibir. Como explica

tribunais superiores. **Revista de Direito Civil Contemporâneo**. v. 31. ano 9. p. 129-153. São Paulo: Ed. RT, abr./jun. 2022. p. 12. Disponível em: <http://revistadostribunais.com.br/maf/app/document?stid=st-rql&marg=DTR-2022-9860>. Acesso em: 20/07/2022.)

⁵⁸⁴ MONTEIRO, Artur Pericles Lima; CRUZ, Francisco Brito; SILVEIRA, Juliana Fonteles da; VALENTE, Mariana Giorgetti. **Armadilhas e caminhos na regulação da moderação de conteúdo** - Diagnósticos & Recomendações. São Paulo: InternetLab, 2021. p. 8.

⁵⁸⁵ Monteiro, Cruz, Silveira e Valente, explicam que isso corroeu, poque a “jurisprudência brasileira tendia a admitir que esses provedores fossem responsabilizados pelo conteúdo postado por usuários sempre que, tendo apenas uma notificação requerendo sua remoção, mantivessem-no disponível. Esse regime de responsabilização de intermediários, conhecido como notice and takedown, levantava preocupações por criar um incentivo às plataformas para realizar a remoção a despeito da procedência da queixa, considerando o risco que assumiriam caso deixassem o conteúdo no ar.” (MONTEIRO, Artur Pericles Lima; CRUZ, Francisco Brito; SILVEIRA, Juliana Fonteles da; VALENTE, Mariana Giorgetti. **Armadilhas e caminhos na regulação da moderação de conteúdo** - Diagnósticos & Recomendações. São Paulo: InternetLab, 2021. p. 8.)

Nitrini, “a partir da juridicidade do conteúdo: tratando-se de conteúdo ilícito, pode determinar sua derrubada nos termos do Marco Civil da Internet.”⁵⁸⁶

O que apontamos como temerário é fechar os olhos, totalmente, à moderação de conteúdo realizada pelas próprias plataformas *online*. Deixar esse árduo processo nas mãos de um poder judiciário, sobrecarregado, burocrático e lento, não condiz com a fluidez e velocidade da produção de conteúdo em ambientes *online*, e o quão ágil os moderadores devem ser para manter um ambiente minimamente saudável, aberto e democrático. Importante, por outro lado, consignar interpretação de Keller, no sentido de que pretensão da lei é regular expressamente, a responsabilidade de intermediários nos casos de conteúdos não removidos após ordem judicial. Com isso, o Marco Civil não se ocupa em controlar a moderação de conteúdo, e, por isso, não apresenta nenhum óbice aos provedores restringirem conteúdos postados por usuários.⁵⁸⁷

Nitrini argumenta sobre esse posicionamento – da prerrogativa exclusiva do Poder Judiciário para decidir sobre a remoção de conteúdo –, “briga com realidade: ignora, no caso das grandes redes sociais, o que elas são e o que fazem.”⁵⁸⁸

Nesse mesmo raciocínio, Goldman, reconhece a qualidade e legitimidade dos tribunais pelo alto grau de especialização, precisão na admissão de provas e na aplicação da lei a elas, adotando procedimentos fomentadores de confiança, sobretudo quanto aos seus resultados.⁵⁸⁹ Entretanto, identifica que os tribunais podem não estar bem-posicionados para determinar soluções para violações *online* porque suas decisões: “(1) demoram muito, com possíveis danos durante a pendência; (2) muitas vezes custam mais do que o valor social da disputa; e (3) pode ser limitado por problemas jurisdicionais que atingem os litigantes.”⁵⁹⁰

Por outro lado, como defendemos neste estudo, os remédios alternativos devem ser priorizados em detrimento, inclusive, de decisões judiciais. Acrescente-se, mesmo assim, a existência de falhas e equívocos causados pelas plataformas, as quais, muitas vezes, ao aplicar

⁵⁸⁶ NITRINI, Rodrigo Vidal. **Liberdade de expressão nas redes sociais: o problema da remoção de conteúdos pelas plataformas**. Belo Horizonte: Editora Dialética, 2021. p. 165.

⁵⁸⁷ KELLER, Clara Iglesias. Policy by judicialisation: the institutional framework for intermediary liability in Brazil. **International Review of Law, Computers & Technology**, p. 1–19, 2020. Disponível em: doi.org/10.1080/13600869.2020.1792035. Acesso em: 01/08/2022.

⁵⁸⁸ NITRINI, Rodrigo Vidal. **Liberdade de expressão nas redes sociais: o problema da remoção de conteúdos pelas plataformas**. Belo Horizonte: Editora Dialética, 2021. p. 163.

⁵⁸⁹ GOLDMAN, Eric. Content Moderation Remedies. **Michigan Technology Law Review**. v. 28, n. 1. p. 1-59, 2021. p. 52. Disponível em: <https://repository.law.umich.edu/mtlr/vol28/iss1/2>. Acesso em: 02/11/2021.

⁵⁹⁰ Tradução livre do original: “(1) take a long time, with harm possibly accruing during the pendency; (2) often cost more than the social value of the dispute; and (3) may be limited by jurisdictional problems reaching the disputants.” (GOLDMAN, Eric. Content Moderation Remedies. **Michigan Technology Law Review**. v. 28, n. 1. p. 1-59, 2021. p. 52. Disponível em: <https://repository.law.umich.edu/mtlr/vol28/iss1/2>. Acesso em: 02/11/2021).

os remédios para violações de usuários, evitam a maioria das limitações dos tribunais, criando desafios diferentes; negligenciando o devido processo legal, sem ouvir todos os envolvidos ou escolher remédios que melhor atendam seus interesses, não os dos usuários.

Ainda assim, são preferíveis os remédios aplicados pelos intermediários, apesar dessas limitações, pois estão melhor situados para compreender os principais anseios das comunidades.⁵⁹¹ Bem como acrescenta Goldmen, “eles também enfrentam consequências de mercado por calibrarem incorretamente seus remédios. Por essas razões, na maioria das situações, devemos preferir que os serviços de *Internet*, não tribunais, decidem os remédios apropriados.”⁵⁹²

Destarte, pontuamos a conformação dessa dicotomia: de um lado, os mecanismos tradicionais do direito do estado, materializado no Marco Civil da Internet no Brasil; e, do outro, os remédios alternativos de moderação de conteúdo, os quais são aplicados por empresas de mídias sociais por meio dos arranjos de governança privada de moderação de conteúdo. Deve-se, portanto, haver o exercício, pelo judiciário da sua competência vigente para processar e julgar os pedidos de remoção de conteúdos, tal como previsto no Art. 19 do Marco Civil da Internet, segundo o juízo de legalidade ou ilegalidade. As plataformas *online*, nesse caso, exercem sua governança privada de moderação de conteúdo em conformidade com suas normas próprias.

Cabendo, à vista disso, ao ordenamento jurídico brasileiro, o reconhecimento da autonomia para tomada de decisões pelas plataformas *online*⁵⁹³, com a sobreposição de direitos fundamentais nas relações entre particulares, especialmente, à liberdade das empresas definirem regras de discursos em seus ambientes, observando-se a liberdade de expressão dos usuários. Destarte, se as redes sociais não possuírem essa liberdade, como expõe Nitrini, isso

⁵⁹¹ Nitrini complementa, que “as redes sociais podem tomar medidas menos drásticas (ou seja, com menor grau de intervenção ao direito de liberdade de expressão) do que apenas a derrubada de postagens, pois há uma gradação de respostas possíveis para conteúdos considerados fronteiriços.” (NITRINI, Rodrigo Vidal. **Liberdade de expressão nas redes sociais**: o problema da remoção de conteúdos pelas plataformas. Belo Horizonte: Editora Dialética, 2021. p. 164).

⁵⁹² Tradução livre do original: “They also face marketplace consequences for miscalibrating their remedies. For these reasons, in most situations, we should prefer that Internet services, not courts, decide the appropriate remedies.” (GOLDMAN, Eric. Content Moderation Remedies. **Michigan Technology Law Review**. v. 28, n. 1. p. 1-59, 2021. p. 53. Disponível em: <https://repository.law.umich.edu/mtlr/vol28/iss1/2>. Acesso em: 02/11/2021).

⁵⁹³ Importante argumento de Virgílio Afonso da Silva, quanto a possibilidade de entes privados promover restrições além das necessárias aos direitos fundamentais nas relações privadas: “exigir que os particulares adotem, nos casos de restrição a direitos fundamentais, apenas as medidas estritamente necessárias – ou seja, as menos gravosas – para o atingimento dos fins perseguidos nada mais é do que retirar-lhes a autonomia de livremente dispor sobre os termos de seus contratos [...]” (SILVA, Virgílio Afonso da. **A constitucionalização do direito**: os direitos fundamentais nas relações entre particulares. Malheiros, 2005. p. 163-164).

implicaria “negar a ideia de que tenham qualquer tipo de autonomia para definir o que é dito em seus ambientes.”⁵⁹⁴

Por conseguinte, dentro da possibilidade de as plataformas tomarem suas decisões no bojo da moderação de conteúdo, estarão, mesmo assim, sujeitas à análise pelo Judiciário, o qual poderá decidir se esta restrição se encontra de acordo com a licitude ou constitucionalidade; bem como examinar a razoabilidade da medida de restrição, independentemente dos critérios adotados pelas plataformas não coincidirem com modelos de legalidade ou ilegalidade. Essa decisão do judiciário é típica dos casos envoltos à eficácia horizontal de direito fundamentais, nos quais há pressuposto de uma dimensão de autonomia entre as partes privadas.⁵⁹⁵ Inclusive, essa possibilidade estimula as plataformas apresentarem fundamentos mais claros, fortalecendo a transparência no processo de moderação de conteúdo.

Nesse panorama e com o afã de aperfeiçoar essa disposição legislativa, Nitrini propõe uma atualização do texto do Art. 19 do Marco Civil da Internet, para que:

- a) trabalhe prioritariamente com obrigações e incentivos para que a governança de conteúdos por redes sociais seja pautada por uma lógica de transparência, prestação de contas e implementação de um devido processo digital;
- b) mantenha a regra geral de isenção de responsabilidade civil a intermediários digitais, com exceções limitadas a situações específicas e justificadas.⁵⁹⁶

Isto posto, para construir o melhor cenário brasileiro quanto à responsabilidade de intermediários na *internet*, é imperativa a adoção das perspectivas dos eixos dos Princípios de Santa Clara, principalmente sobre a perspectiva procedimental,⁵⁹⁷ equacionando dois pilares importantes para uma moderação de conteúdo aceitável: primeiro, a fixação de direitos como uma forma de ajustar a relação entre plataformas e seus usuários, proporcionando maior

⁵⁹⁴ NITRINI, Rodrigo Vidal. **Liberdade de expressão nas redes sociais**: o problema da remoção de conteúdos pelas plataformas. Belo Horizonte: Editora Dialética, 2021. p. 168.

⁵⁹⁵ NITRINI, Rodrigo Vidal. **Liberdade de expressão nas redes sociais**: o problema da remoção de conteúdos pelas plataformas. Belo Horizonte: Editora Dialética, 2021. p. 171-172.

⁵⁹⁶ NITRINI, Rodrigo Vidal. **Liberdade de expressão nas redes sociais**: o problema da remoção de conteúdos pelas plataformas. Belo Horizonte: Editora Dialética, 2021. p. 180.

⁵⁹⁷ Termo utilizado por Monteiro, Cruz, Silveira e Valente, que argumentam: “Como propostas regulatórias podem equacionar essas duas chaves? Uma abordagem que pode ser promissora parte do que chamamos de uma perspectiva procedimental. Essa abordagem seria uma forma de disciplinar a atuação das plataformas quando elas agirem sobre seus usuários, sem, contudo, engessá-la. No lugar de tachar qualquer moderação como censura, essa perspectiva busca regular esse processo, reconhecendo direitos aos usuários e proporcionando mais transparência, tanto aos usuários envolvidos numa controvérsia quanto à sociedade em geral.” (MONTEIRO, Artur Pericles Lima; CRUZ, Francisco Brito; SILVEIRA, Juliana Fonteles da; VALENTE, Mariana Giorgetti. **Armadilhas e caminhos na regulação da moderação de conteúdo** - Diagnósticos & Recomendações. São Paulo: InternetLab, 2021. p. 28.)

clareza e transparência, a fim de facilitar a compreensão dos indivíduos sobre os motivos de determinadas medidas; segundo a busca por um ambiente digital aberto e democrático.⁵⁹⁸

3.4.2. Proteção de Dados

É cediço o debate sobre a proteção de dados e privacidade na *internet* ter ocorrido antes do Marco Civil da Internet⁵⁹⁹, mesmo assim a regulação digital no Brasil é bem recente e como visto no item anterior, o Marco Civil da Internet, estabelece a proteção de dados no ambiente digital, contudo com a presença de lacunas legislativas.

Nesse contexto, verifica-se que a Constituição Federal de 1988, estipula no Art. 5º, XII, o sigilo das comunicações de dados, embora não fixe expressamente um direito fundamental à proteção e livre disposição dos dados pelo próprio usuário. Por outro lado, a Carta Magna confere proteção de dados pessoais, ainda que de forma indireta, com a previsão da ação de *habeas data*, no Art. 5º, LXXII.⁶⁰⁰ Tal ação constitui um remédio constitucional visando assegurar ao indivíduo o conhecimento e a possibilidade de buscar a retificação de dados constantes de registros ou bancos de dados de entidades governamentais ou de caráter público, além de constituir uma garantia procedimental do exercício da autodeterminação informacional.⁶⁰¹

Recentemente, fora promulgada a Emenda Constitucional n. 115, de 2022, reconhecendo explicitamente, como um direito fundamental, a proteção de dados pessoais no catálogo constitucional de direitos, mediante a inclusão no Art. 5º da Constituição Federal, o inciso LXXIX, com a seguinte redação: “é assegurado, nos termos da lei, o direito à proteção dos dados pessoais, inclusive nos meios digitais.” Bem como, inseriu o inciso XXVI, no Art. 21, da CF: “organizar e fiscalizar a proteção e o tratamento de dados pessoais, nos termos da lei”, e o inciso XXX, no Art. 22, da CF: “proteção e tratamento de dados pessoais”.

⁵⁹⁸ Esse pensamento, está alicerçado não somente nos princípios de Santa Clara, como também no recente relatório “Olhando Al Sur”, Al Sur é um consórcio de organizações e acadêmicos que procuram reforçar os direitos humanos no ambiente digital na América Latina. Conferir: CAMPO, Agustina Del; SCHATZKY, Morena; HERNÁNDER, Laura. LARA, J. Carlos. **Olhando Al Sur**. Rumo a novos consensos regionais em matéria de responsabilidade de intermediários na Internet. 2021. Disponível em: <https://www.alsur.lat/sites/default/files/2021-06/Responsabilidad%20de%20intermediarios%20PT.pdf>. Acesso em: 02/08/2022.

⁵⁹⁹ POLETTO, Álerton Emanuel. **O dever jurídico de moderação de conteúdo pelo Facebook**. São Paulo: Editora Dialética, 2022. p. 84-85.

⁶⁰⁰ SARLET, Ingo Wolfgang; SAAVEDRA, Giovanni Agostini. Fundamentos Jusfilosóficos e Âmbito de Proteção de Direito Fundamental à Proteção de Dados Pessoais. **Revista de Direito Público**. Volume 17, n. 93, p. 33-57, Brasília, 2020. p. 41-42. Disponível em: <https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/4315>. Acesso em: 24/10/2021.

⁶⁰¹ MENDES, Laura Schertel. *Habeas data* e autodeterminação informativa: dois lados da mesma moeda. **Revista Direitos Fundamentais & Justiça**, Ano. 12, n. 39, p. 185-216, jul./dez. 2018.

Antes da previsão constitucional, a doutrina já vinha declinando que o direito à proteção de dados pessoais encontrava tutela constitucional na Constituição Federal de 1988 de forma indireta, aduzindo que se associava aos princípios e direitos fundamentais de caráter geral e especial, como o princípio da dignidade da pessoa humana, direitos à privacidade, à intimidade, ao livre desenvolvimento da personalidade e o direito geral de liberdade.⁶⁰²

Neste trabalho, não será objeto de análise profunda sobre a consagração textual, como um direito fundamental autônomo, a proteção de dados. Sublinhamos, contudo, a importância dessa positivação formal, gerando uma articulação e interação com outros direitos, passando a ter um regime jurídico-constitucional quanto ao seu perfil de direito fundamental material e formal, consagrado no texto constitucional, bem como na doutrina e jurisprudência brasileira.

No âmbito infraconstitucional, foi editada a Lei n. 13.709 de agosto de 2018, a Lei Geral de Proteção de Dados – LGPD, dispondo sobre o tratamento de dados pessoais, buscando proteger os direitos fundamentais de liberdade, privacidade e livre desenvolvimento da personalidade, tanto nos ambientes digitais, como fora deles. Conforme se extrai do art. 1º, vejamos:

Art. 1º Esta Lei dispõe sobre o tratamento de dados pessoais, inclusive nos meios digitais, por pessoa natural ou por pessoa jurídica de direito público ou privado, com o objetivo de proteger os direitos fundamentais de liberdade e de privacidade e o livre desenvolvimento da personalidade da pessoa natural.

Observa-se, o fato dos titulares do direito à proteção de dados serem as pessoas naturais, devendo, a nosso ver, ser coadunado com a previsão do Art. 5º, *caput*, da CF, onde prevê que são titulares dos direitos fundamentais os brasileiros e estrangeiros residentes no país. Atualmente, segue sendo majoritário o entendimento, de a proteção de dados pessoais possuir como titulares apenas pessoas naturais, por não se compatibilizar com as pessoas jurídicas. Reconhece-se, que as pessoas jurídicas já são protegidas, inclusive na perspectiva jusfundamental, por outros direitos e garantias – como o sigilo industrial e comercial, propriedade imaterial. Contudo, como aduz Sarlet é “questionável que a inclusão das pessoas jurídicas no polo subjetivo ativo dos direitos à privacidade e intimidade, bem como do direito à proteção de dados pessoais, implique ganho real qualitativo de proteção.”⁶⁰³

Tem-se como destinatários do direito estabelecido na Lei, tanto o Estado como os particulares. Nesse ponto, caberá aos autores estatais velar pela aplicação e concretização do

⁶⁰² SARLET, Ingo Wolfgang. Proteção de Dados Pessoais como Direito Fundamental na Constituição Brasileira de 1988: contributo para a construção de uma dogmática constitucionalmente adequada. **Direitos Fundamentais & Justiça**. Ano 14, n. 42, p. 179-218, jan./jun. Belo Horizonte, 2020. p. 184.

⁶⁰³ SARLET, Ingo Wolfgang. Proteção de Dados Pessoais como Direito Fundamental na Constituição Brasileira de 1988: contributo para a construção de uma dogmática constitucionalmente adequada. **Direitos Fundamentais & Justiça**. Ano 14, n. 42, p. 179-218, jan./jun. Belo Horizonte, 2020. p.204.

direito à proteção de dados, primando pela sua máxima eficácia e efetividade concreta. Nesse ponto, cumpre lançar luz quanto ao objeto deste trabalho: a necessidade de atores privados serem sujeitos e convocados a aplicarem e concretizarem o direito à proteção de dados.

Tal como visto no Capítulo II, é amplamente reconhecido no Brasil a eficácia dos direitos fundamentais nas relações privadas, existindo divergência quanto ao modo que se materializa essa vinculação e eficácia, Sarlet leciona que se dá de modo diferenciado, numa perspectiva multinível (por passar tanto no campo do direito internacional público, como no direito pátrio), e reconhece:

a proteção dos dados pessoais, seja em que contexto for, mas em especial no ambiente digital, não se pode admitir uma esfera de atuação privada completamente livre dos direitos fundamentais, gerando uma espécie de imunidade, tanto mais perigosa – no que concerne a violações de direitos – quanto mais força tiverem os atores privados que operam nesse cenário. Por isso, um controle rigoroso das restrições a direitos fundamentais na esfera das relações privadas, inclusive em caráter preventivo, levando em conta os deveres de proteção estatais também em face de perigos e riscos, é de ser levado a efeito (inclusive!) pelos Tribunais. Tal controle, contudo, deve levar a sério, em primeira em primeira linha, as opções legislativas, mas ao mesmo tempo, não hesitar quando se trata de reconhecer e declarar eventual inconstitucionalidade, pois, do contrário, a proteção dos dados pessoais poderá estar comprometida.⁶⁰⁴

Um outro aspecto digno de nota, diz respeito à limitação deste direito fundamental, sendo, tal como os demais, também submetido a limites e passível de restrições, com objetivo de proteger outros direitos fundamentais ou bens jurídicos de estrutura constitucional. Ilustramos, com uma jurisprudência do Supremo Tribunal Federal, que ao julgar o SS n. 3.902, de Relatoria do Ministro Teori Zavascki (julgado em 23/04/2015), envolvendo uma lide entre o direito de acesso a informações de caráter público em poder de órgãos públicos e o direito à proteção de dados pessoais sensíveis dos servidores públicos. A Suprema Corte, entendeu a proteção da privacidade dos servidores públicos menor em relação à do cidadão comum, de modo a considerar constitucionalmente legítima e proporcional a divulgação nominal e individualizada dos seus vencimentos e benefícios.⁶⁰⁵

Ingo W. Sarlet expõe que, o direito fundamental à proteção de dados possui dupla dimensão subjetiva e objetiva. Na condição subjetiva, assevera o direito à proteção de dados pessoais se decodificando em um “conjunto heterogêneo de posições subjetivas de natureza defensiva (negativa), mas também assume a condição de direito a prestações, cujo objeto

⁶⁰⁴ SARLET, Ingo Wolfgang. Proteção de Dados Pessoais como Direito Fundamental na Constituição Brasileira de 1988: contributo para a construção de uma dogmática constitucionalmente adequada. **Direitos Fundamentais & Justiça**. Ano 14, n. 42, p. 179-218, jan./jun. Belo Horizonte, 2020. p.210.

⁶⁰⁵ SUPREMO TRIBUNAL FEDERAL. **Suspensão de Segurança 3.902/SP**. Rel. Min. Teori Zavascki. 2015.

consiste em uma atuação do estado mediante a disponibilização de prestações de natureza fática ou normativa”.⁶⁰⁶

Nessa perspectiva, o enunciado nos arts. 17 e 18 da LGPD, constata-se posições jurídicas subjetivas (direitos) atribuídas ao titular dos dados pessoais objeto da proteção legal, concretizando e delimitando o próprio âmbito de proteção do direito fundamental à proteção de dados, e guardando certa aparência com o rol de posições jurídico-constitucionais diretamente e habitualmente associadas à dupla função de tal direito como direito negativo (defesa) e positivo (a prestações).⁶⁰⁷ Vejamos a transcrição do catálogo legal referido:

Art. 17. Toda pessoa natural tem assegurada a titularidade de seus dados pessoais e garantidos os direitos fundamentais de liberdade, intimidade e de privacidade, nos termos desta lei.

Art. 18. O titular dos dados pessoais tem direito a obter do controlador, em relação aos dados do titular por ele tratados, a qualquer momento e mediante requisição:

I - confirmação da existência de tratamento;

II - acesso aos dados;

III - correção de dados incompletos, inexatos ou desatualizados;

IV - anonimização, bloqueio ou eliminação de dados desnecessários, excessivos ou tratados em desconformidade com o disposto nesta Lei;

V - portabilidade dos dados a outro fornecedor de serviço ou produto, mediante requisição expressa, de acordo com a regulamentação da autoridade nacional, observados os segredos comercial e industrial;

VI - eliminação dos dados pessoais tratados com o consentimento do titular, exceto nas hipóteses previstas no art. 16 desta Lei;

VII - informação das entidades públicas e privadas com as quais o controlador realizou uso compartilhado de dados;

VIII - informação sobre a possibilidade de não fornecer consentimento e sobre as consequências da negativa;

IX - revogação do consentimento, nos termos do §5º do art. 8º desta Lei. [...]

Art. 20. O titular dos dados tem direito a solicitar a revisão de decisões tomadas unicamente com base em tratamento automatizado de dados pessoais que afetem seus interesses, incluídas as decisões destinadas a definir o seu perfil pessoal, profissional, de consumo e de crédito ou os aspectos de sua personalidade.

Art. 21. Os dados pessoais referentes ao exercício regular de direitos pelo titular não podem ser utilizados em seu prejuízo.

Art. 22. A defesa dos interesses e dos direitos dos titulares de dados poderá ser exercida em Juízo, individual ou coletivamente, na forma do disposto na legislação pertinente, acerca dos instrumentos de tutela individual e coletiva.

Por não se tratar de um rol taxativo, sendo cabível outras possibilidade, mesmo não previstas expressamente na Constituição ou em legislação infraconstitucional, percebe-se uma

⁶⁰⁶ SARLET, Ingo Wolfgang. Proteção de Dados Pessoais como Direito Fundamental na Constituição Brasileira de 1988: contributo para a construção de uma dogmática constitucionalmente adequada. **Direitos Fundamentais & Justiça**. Ano 14, n. 42, p. 179-218, jan./jun. Belo Horizonte, 2020. p. 194.

⁶⁰⁷ SARLET, Ingo Wolfgang. Proteção de Dados Pessoais como Direito Fundamental na Constituição Brasileira de 1988: contributo para a construção de uma dogmática constitucionalmente adequada. **Direitos Fundamentais & Justiça**. Ano 14, n. 42, p. 179-218, jan./jun. Belo Horizonte, 2020. p. 195-196.

característica, não muito rara no direito brasileiro, que é a reprodução de direitos já expressos no ordenamento jurídico, como se vê dos direitos de liberdade, intimidade e privacidade esculpido no Art. 17, bem como a previsão do acesso à Justiça, expresso no Art. 22.

3.4.3. Moderação de Conteúdo e Período Eleitoral

Nos últimos anos, sobretudo desde as Eleições presidenciais de 2018, as notícias falsas ocuparam lugar de proeminência no debate público, transcendendo o objetivo de, meramente, angariar votos e, conseqüentemente, alcançar resultados eleitorais.⁶⁰⁸ Passaram a influir em grandes eventos importante na sociedade, como se constatou durante a pandemia causada pelo COVID-19.⁶⁰⁹

No Brasil, de forma especial, verifica-se uma campanha maciça de desinformação,⁶¹⁰ durante o período eleitoral, posto que, durante esse período se destaca os posicionamentos políticos, os quais influenciam diretamente na decisão popular, podendo interferir nos resultados do pleito eleitoral. Assim, a campanha de desinformação é retroalimentada por interesses muitas vezes não republicanos, considerando que é necessário o financiamento para colocar em funcionamento o sistema, visando como retorno a conquista de cargos públicos.

Desse raciocínio, extrai-se que um dos principais pontos a ser combatido, durante o período eleitoral no Brasil, seria a disseminação de desinformação, logicamente respeitando os limites constitucionais regulatórios e, nesse sentido, a Justiça Eleitoral tem empreendido esforços, em sua função fiscalizatória, quanto ao combate das propagandas eleitorais

⁶⁰⁸ Nesse ponto, calha transcrever as observações alinhavadas por Sarlet e Siqueira: “Dentre algumas medidas adotadas pelo TSE, para as últimas eleições ocorridas em 2018, destaca-se a criação do Conselho Consultivo sobre Internet e Eleições, cujas atribuições estão atreladas ao desenvolvimento de pesquisas da correta aplicação das normas eleitorais, à emissão de pareceres sobre matérias que lhe sejam submetidas pela Presidência do TSE, e à proposição de ações para o aperfeiçoamento dessas normas. Além disso, foi assinado um termo de compromisso, proposto em junho de 2018, com 31 partidos políticos, os quais se comprometeram em colaborar contra o compartilhamento de informações falsas nas eleições ocorridas naquele ano. Ainda que não conste no rol de partidos, PSTU afirma ter assinado o termo. Ademais, foi também assinado um memorando de entendimento entre a Justiça Eleitoral e o Conselho de Política Institucional e as mídias sociais para coibir a disseminação das então chamadas *fake news*, com o Facebook e com o Google, no fim de junho de 2018”. (SARLET, Ingo Wolfgang; SIQUEIRA, Andressa de Bittencourt. Liberdade de expressão e seus limites numa democracia: o caso das assim chamadas “fake news” nas redes sociais em período eleitoral no Brasil. **Revista de Estudos Institucionais**. vol. 6, n. 2, p. 534-578, maio/ago, 2020. p. 542)

⁶⁰⁹ “Uma importante pesquisa debate a ideia de que a partir do momento que a Organização Mundial da Saúde (OMS), decreta a pandemia, o mundo passa a conviver de maneira íntima com a chamada infodemia, que nada mais é que a disseminação de notícias falsas que tem como intuito desinformar as pessoas e agravar ainda mais a saúde mental e psicológica das pessoas”. (DOS SANTOS, Regina Consolação et al. Fake news e a desinformação sobre COVID-19: revisão integrativa de literatura. **Research, Society and Development**, v. 11, n. 10, 2022. Disponível em: <https://rsdjournal.org/index.php/rsd/article/view/33124>. Acesso em: 25/11/2022.)

⁶¹⁰ Quanto a denominação que melhor expressa a complexidade do fenômeno, entendemos ser a desinformação (popularizada como a expressão *fake news*) a mais apropriada, pois abrange situações potencialmente típicas. Ademais, é o termo adotado pelo próprio Tribunal Superior Eleitoral (TSE) na Resolução 23.610/2019.

irregulares durante o período eleitoral no âmbito virtual. Demonstraremos, de forma sucinta, um panorama da atuação da Justiça Eleitoral, sobretudo quanto ao combate à desinformação no período eleitoral.

a) *Checagem de informações*: com a transferência do debate público e político para o ambiente digital, tem-se percebido que a legislação eleitoral e os atos normativos do Tribunal Superior Eleitoral – TSE, não acompanharam, a contento, essa evolução – o que de todo modo é perfeitamente compreensível – a simples aplicação de dispositivos até então endereçados à mídia tradicional (rádio, TV, jornais impressos etc.) não se mostraram adequadas ao ambiente virtual. Como exemplo, a aplicação do direito de resposta tem se mostrado ineficaz nas plataformas digitais, dada a velocidade das informações e abrangência, não sendo oferecido ao candidato alvo de informações falsas, a chance de reposta ao público.

Nesse sentido, o TSE procurou abrir um diálogo institucional com empresas de redes sociais, estimulando-as a adotarem, em suas plataformas, condutas de autocontenção da desinformação, em especial, no período eleitoral⁶¹¹. Desde 2018, a Justiça Eleitoral atua no combate à disseminação de notícias falsas relacionadas às eleições e à própria Justiça Eleitoral, sendo lançado, em 2019, o Programa de Enfrentamento à Desinformação, tornando-se permanente a partir de 2021. Sendo realizada o firmamento de parceria com as principais redes sociais e plataformas digitais de compartilhamento de mensagens e vídeos, onde foram destacados os principais pontos, dentre eles: “os perigos da proliferação de notícias falsas para a estabilidade democrática, especialmente no contexto de um pleito geral, e a necessidade da cooperação das plataformas digitais nas medidas que visem coibir ou neutralizar a divulgação de conteúdo inautêntico pela *internet*.”⁶¹²

Ademais, tem-se desenvolvido, junto às plataformas digitais, na construção de mecanismos inibidores da proliferação de desinformação, com desenvolvimento de filtros e estrutura de vigilância, como ferramentas de *fact checking* (checagem de informação). Como explicam Batini e Oliveira, essa checagem deve passar por duas etapas: “uma, com a análise dos elementos da notícia, e outra, com a verificação do conteúdo em fontes seguras de

⁶¹¹ TSE reúne-se com Google, Facebook, WhatsApp e Twitter para alinhar estratégias de combate à desinformação nas Eleições 2020. (TRIBUNAL SUPERIOR ELEITORAL. Eleições 2022: TSE reúne-se com Google, Facebook, WhatsApp e Twitter para alinhar estratégias de combate à desinformação nas Eleições 2020. tse.jus.br. Brasília, 12 de nov. 2019. Disponível em: <https://www.tse.jus.br/comunicacao/noticias/2019/Novembro/tse-se-reune-com-google-facebook-whatsapp-e-twitter-para-alinhar-estrategias-de-combate-a-desinformacao-nas-eleicoes-2020>. Acesso em: 30/11/2022).

⁶¹² TRIBUNAL SUPERIOR ELEITORAL. Eleições 2022: TSE firma parceria com Facebook para combate à desinformação. tse.jus.br. Brasília, 11 de fev. de 2022. Disponível em: <https://www.tse.jus.br/comunicacao/noticias/2022/Fevereiro/eleicoes-2022-tse-firma-parceria-com-facebook-para-combate-a-desinformacao>. Acesso em: 30/11/2022.

informação.”⁶¹³ Aduzem, ainda, que após a primeira etapa, “são feitas consultas em sites de grandes meios de comunicação e fontes oficiais relacionadas ao conteúdo da notícia para apurar se ela é realmente verdadeira.”⁶¹⁴

b) *Controle da propaganda*: a desinformação pode configurar propaganda irregular, que segue a disposição da Lei n. 9.504/97, nos Artigos 57-A a 57-J, além das Resoluções do TSE e, no intuito de enquadrar a desinformação como propaganda irregular, o Artigo 242 do Código Eleitoral, dispõe qualquer forma ou modalidade de propaganda, deve mencionar a legenda do partido, não podendo empregar meios publicitários destinados a criar, artificialmente, na opinião pública, estados mentais, emocionais ou passionais. Desta forma, tem-se a desinformação sendo a geradora de sentimentos fortes, não raramente como ódio, medo, revolta, devendo ser considerada como irregular, a propaganda eleitoral, que dissemina de forma orquestrada, postagens em período eleitoral capaz de perturbar a normalidade da construção da convicção do eleitor.

c) *Balizas jurisprudenciais*: a crítica no período eleitoral é absolutamente normal e faz parte do debate político eleitoral, integrando típico discurso de opositores.⁶¹⁵ Desta forma, é assegurado a livre manifestação do pensamento, inclusive no ambiente digital, somente devendo sofrer restrições quando houve ofensa à honra de terceiros ou divulgação de fatos sabidamente inverídicos.⁶¹⁶

Nesse passo, para caracterizar uma mensagem/publicação/postagem como sabidamente inverídica, segundo a jurisprudência, “deve conter inverdade flagrante que não apresente controvérsias. Não é possível transformar o pedido de resposta em processo investigatório com intuito de comprovar a veracidade das versões controversas sustentadas pelas partes.”⁶¹⁷ O Tribunal Superior Eleitoral tem deitado entendimento quanto ao exercício do direito de resposta, em prol da liberdade de expressão, somente devendo ser concedido de forma excepcional, permitindo-o, apenas quando for possível extrair, da afirmação apontada como sabidamente inverídica, ofensa de caráter pessoal a candidato, partido ou coligação. Senão vejamos:

[...] O direito de resposta não se presta a rebater a liberdade de expressão e de opinião que são inerentes à crítica política e ao debate eleitoral. [...]. O fato

⁶¹³ BATINI, Silvana; OLIVEIRA, Neide M. C. Cardoso de. Desinformação e Eleições 2020: caminhos possíveis. **Revista Justiça Eleitoral em Debate**. v. 10. n. 1. Rio de Janeiro, 2020.

⁶¹⁴ BATINI, Silvana; OLIVEIRA, Neide M. C. Cardoso de. Desinformação e Eleições 2020: caminhos possíveis. **Revista Justiça Eleitoral em Debate**. v. 10. n. 1. Rio de Janeiro, 2020.

⁶¹⁵ Nesse sentido, é a jurisprudência do TSE: AC nº 349/2002, Rel. Min. Sálvio de Figueiredo; AC. nº 588/2002, rel. Min. Caputo Bastos

⁶¹⁶ Recurso Especial Eleitoral nº 204014, Acórdão, Relator(a) Min. Luciana Christina Guimarães Lóssio, Publicação: DJE - Diário de justiça eletrônico, Data 10/11/2015,

⁶¹⁷ Representação nº 367516 – Rel. Henrique Neves – j. 26.10.2010

sabidamente inverídico a que se refere o art. 58 da Lei nº 9.504/97, para fins de concessão de direito de resposta, é aquele que não demanda investigação, ou seja, deve ser perceptível de plano. [...]. (Representação nº 139448, Acórdão, Relator(a) Min. Admar Gonzaga Neto, Publicação: PSESS - Publicado em Sessão, Data 02/10/2014).

Dos julgados realizados, com referência às Eleições de 2018, constata-se o TSE adotando uma postura de conceder máxima obediência à liberdade de expressão, procurando restringir somente postagens mais duras, críticas ou mesmo satíricas contra adversários políticos, atuando de forma mais enérgica, com a remoção de conteúdos das redes, somente quando constatada presença de uma mentira flagrante, com tons de ofensa à honra ou à imagem de um dos *players* eleitorais.⁶¹⁸

Contudo, sobretudo, durante o segundo turno das Eleições de 2018 houve um acirramento e consequente aumento de tensão no cenário eleitoral, transbordando o debate político-eleitoral, sendo marcado com a disseminação de *fake news*, agora não mais quanto aos candidatos, partidos ou coligações, mas sim ataques à instituição da Justiça Eleitoral e ao sistema eletrônico de votação, na figura da utilização da urna eletrônica. Com isso, o TSE teve de adotar postura mais dura, para garantir a higidez do processo eleitoral e da própria democracia.⁶¹⁹

Verifica-se, como principal julgado, representando essa virada no posicionamento do TSE, a Representação Eleitoral nº 060129842, na qual a Corte Eleitoral determinou ao Google Brasil e ao Facebook Serviços On-Line do Brasil que excluíssem um vídeo que circulava em páginas *internet*, no qual constava o então candidato a presidente da República Jair Bolsonaro, tecendo críticas às urnas eletrônicas.

A partir desse julgado, a Justiça Eleitoral, passou a incorporar um discurso de defesa da própria instituição, especificamente, do sistema eletrônico de votação, moderando discursos que tentassem desacreditar o sistema de votação, em defesa do processo eleitoral. Com isso, inaugurava-se uma nova fase na atuação do TSE, quanto ao tema das notícias fraudulentas (*fake news*) ou da desinformação.

⁶¹⁸ ALARCON, Anderson de Oliveira. **Sociedade da (des)informação**: análise da jurisprudência e das decisões do TSE sobre fake news nas eleições de 2018 e de 2020. 2022. Tese de Doutorado – Universidade Federal do Paraná, Setor de Ciências Humanas, Programa de Pós-Graduação em Ciência Política. Curitiba, 2022. p. 93.

⁶¹⁹ Alarcon, resume com maestria o que a *fake news* e a desinformação tenta atacar no processo eleitoral: “Primeiro: a institucionalidade eleitoral, as autoridades eleitorais. Deslegitimá-las o máximo possível. Segundo: a campanha. Atacar o adversário, prejudicar a reputação, a credibilidade. Terceiro: o processo eleitoral como um processo. Pois o pipocar de ataques à credibilidade ou fidedignidade do sistema de votação e, portanto, do processo eleitoral como um processo que revela fórmula de constituição legítima da representação popular, acabou desencadeando uma reação mais dura.” (ALARCON, Anderson de Oliveira. **Sociedade da (des)informação**: análise da jurisprudência e das decisões do TSE sobre fake news nas eleições de 2018 e de 2020. 2022. Tese de Doutorado – Universidade Federal do Paraná, Setor de Ciências Humanas, Programa de Pós-Graduação em Ciência Política. Curitiba, 2022. p. 95)

Nessa esteira, e visando o processo de preparação para as Eleições de 2022, o TSE realizou o II Seminário Internacional Desinformação e Eleições, em outubro de 2021. Oportunidade na qual fora realizado um amplo debate sobre os desafios no combate à desinformação, discutindo mecanismos para impedir ou diminuir a divulgação de informações falsas nas eleições de 2022.

Nessa evolução de pensamento, no combate à desinformação, desde as Eleições de 2018, Alarcon resume que a Corte Eleitoral tem retirado as seguintes lições:

a) a cooperação interinstitucional é fundamental, pois os riscos são comuns e, nesse sentido, todos devem se ajudar; b) trabalhar com recursos disponíveis no sentido de aplicar normas e critérios com vistas ao máximo efeito, porém, sem ferir a liberdade de expressão e; c) reconhecer e valorizar a importância da comunicação como área estratégica, que deve andar aliada com a equipe de tecnologia para bem informar.⁶²⁰

Quanto à alteração da postura da Corte Eleitoral, apontamos um caso paradigmático representando essa nova visão contra a disseminação de desinformação, o Recurso Ordinário Eleitoral n.º 0603975-98.2018.6.16.0000 (ROE nº 060397598), envolvendo a cassação do diploma eleitoral outorgado ao Deputado Estadual Fernando Francischini, constituindo o primeiro político com mandato cassado pela disseminação de *fake news* no Brasil.

Em seu voto, o Relator Ministro Luis Felipe Salomão, votou por dar provimento ao recurso, reformando o acórdão local⁶²¹ e julgando procedente a ação, cassando o diploma

⁶²⁰ ALARCON, Anderson de Oliveira. **Sociedade da (des)informação**: análise da jurisprudência e das decisões do TSE sobre fake news nas eleições de 2018 e de 2020. 2022. Tese de Doutorado – Universidade Federal do Paraná, Setor de Ciências Humanas, Programa de Pós-Graduação em Ciência Política. Curitiba, 2022. p. 103.

⁶²¹ Confira, a emenda do TRE/PR: ELEIÇÕES 2018. AÇÃO DE INVESTIGAÇÃO JUDICIAL ELEITORAL. DIVULGAÇÃO DE “LIVE”, EM REDE SOCIAL, FACEBOOK POR DEPUTADO FEDERAL E CANDIDATO A DEPUTADO ESTADUAL, DURANTE O HORÁRIO DA VOTAÇÃO DO PRIMEIRO TURNO, CONTENDO A AFIRMAÇÃO DE QUE ESTAVA COMPROVADO QUE DUAS URNAS ESTAVAM FRAUDADAS OU ADULTERADAS, ALÉM DE OUTRAS CRÍTICAS E CONTEÚDOS FALSOS PROFERIDOS EM TOM SENSACIONALISTA E OFENSIVO. USO INDEVIDO DOS MEIOS DE COMUNICAÇÃO NÃO CONFIGURADO. ABUSO DE AUTORIDADE. CONFIGURAÇÃO. IMUNIDADE PARLAMENTAR MATERIAL QUE NÃO AUTORIZA ATAQUES ÀS INSTITUIÇÕES DEMOCRÁTICAS. PUNIÇÃO EXCEPCIONALMENTE AFASTADA EM VIRTUDE DO ENTENDIMENTO DA MAIORIA DA CORTE DE QUE NÃO HÁ PROVA CABAL DE QUE A CONDUTA DO INVESTIGADO TENHA RESULTADO EM BENEFÍCIO À CANDIDATO NA ELEIÇÃO DE 2018. AÇÃO JULGADA IMPROCEDENTE. 1. Uso indevido dos meios de comunicação não configurado, uma vez que este requer a comprovação da utilização de veículos de imprensa, como rádio, jornal ou televisão, em benefício de determinado candidato, seja pela concessão, em seu favor, de espaço privilegiado na mídia, ou pela crítica abusiva aos demais concorrentes. Aplicativos de mensagens e contas pessoais em redes sociais não se enquadram no conceito legal. 2. Eventual falha de funcionamento das urnas eletrônicas não implica fraude no sistema de votação, inclusive porque fraude exige atuação humana intencional, neste caso, voltada a manipulação do resultado eleitoral. 3. Não se questiona o direito de repassar, por meio das redes sociais, informações sobre relatos de eleitores que estavam enfrentando eventuais problemas técnicos de funcionamento com as urnas eletrônicas, sendo que, inclusive é de se reconhecer a utilidade pública desse tipo de informação e principalmente de quais seriam as medidas que os eleitores poderiam tomar ao se depararem com eventual falha de funcionamento da urna ou qualquer outro problema no momento da votação, de maneira serena e responsável. 4. Um parlamentar não pode propagar irresponsavelmente fatos deturpados, notícias falsas, e teorias conspiratórias sobre fraudes, agentes infiltrados e golpe, com inquestionável potencial de desacreditar

outorgado e o mandato de Francischini, bem como aplicando a sanção de inelegibilidade. O relator Ministro Salomão fora acompanhado por outros cinco ministros: Mauro Campbell, Sérgio Banhos, Edson Fachin, Alexandre de Moraes e Luís Roberto Barroso, os quais entenderam que Francischini praticou crimes ao utilizar seu perfil pessoal no Facebook para promover ataques contra as urnas eletrônicas. O Ministro Carlos Hobach, foi o único voto divergente. Dada a importância e o teor didático, transcrevemos a emenda do julgado:

RECURSO ORDINÁRIO. ELEIÇÕES 2018. DEPUTADO ESTADUAL. AÇÃO DE INVESTIGAÇÃO JUDICIAL ELEITORAL (AIJE). USO INDEVIDO DOS MEIOS DE COMUNICAÇÃO SOCIAL. ABUSO DE PODER POLÍTICO E DE AUTORIDADE. ART. 22 DA LC 64/90. TRANSMISSÃO AO VIVO. REDE SOCIAL. DIA DO PLEITO. HORÁRIO DE VOTAÇÃO. FATOS NOTORIAMENTE INVERÍDICOS. SISTEMA ELETRÔNICO DE VOTAÇÃO. FRAUDES INEXISTENTES EM URNAS ELETRÔNICAS. AUDIÊNCIA DE MILHARES DE PESSOAS. MILHÕES DE COMPARTILHAMENTOS. PROMOÇÃO PESSOAL. IMUNIDADE PARLAMENTAR COMO ESCUDO PARA ATAQUES À DEMOCRACIA. IMPOSSIBILIDADE. GRAVIDADE. CASSAÇÃO DO DIPLOMA. INELEGIBILIDADE. PROVIMENTO.

1. Recurso ordinário interposto pelo Ministério Público contra acórdão prolatado pelo TRE/PR, que, por maioria de votos, julgou improcedente os pedidos em Ação de Investigação Judicial Eleitoral (AIJE) proposta em desfavor de Deputado Estadual eleito pelo Paraná em 2018, afastando o abuso de poder político e o uso indevido dos meios de comunicação social (art. 22 da LC 64/90).

2. Rejeitada a preliminar de inovação recursal aduzida em contrarrazões. Os argumentos contidos no apelo apenas contrapõem a tese da Corte de origem de que a internet e as redes sociais não se enquadram como meios de comunicação.

3. A hipótese cuida de live transmitida ao vivo em rede social, quando em curso a votação no primeiro turno, para mais de 70 mil internautas, e que até 12/11/2018 teve mais de 105 mil comentários, 400 mil compartilhamentos e seis milhões de visualizações. O recorrido – que exercia o cargo de Deputado Federal – noticiou a existência de fraudes em urnas eletrônicas e outros supostos fatos acerca do sistema eletrônico de votação.

4. Sintetizam-se as principais declarações na transmissão: (a) "já identificamos duas urnas que eu digo ou são fraudadas ou adulteradas. [...], eu tô com toda a documentação aqui da própria Justiça Eleitoral"; (b) "nós estamos estourando isso aqui em primeira mão pro Brasil inteiro [...], urnas ou são adulteradas ou fraudadas"; (c) "nosso advogado acabou de confirmar [...], identificou duas urnas que eu digo adulteradas"; (d) "apreensão feita, duas urnas eletrônicas"; (e) "não vamos aceitar que uma empresa da Venezuela, que a tecnologia que a gente não tem acesso, defina a democracia no Brasil"; (f) "só aqui e na Venezuela tem a porcaria da urna eletrônica"; (g) "daqui a pouco nós vamos acompanhar [a apuração dos resultados], sem paradinha técnica, como aconteceu com a Dilma"; (h) "eu uso aqui a minha imunidade parlamentar, que ainda vai até janeiro, independente dessa eleição, pra trazer essa denúncia".

instituições e promover a desordem social. 5. A imunidade parlamentar material não é absoluta e “não confere aos parlamentares o direito de empregar expediente fraudulento, artificioso ou ardiloso, voltado a alterar a verdade da informação, com o fim de desqualificar ou imputar fato desonroso à reputação de terceiros”. Precedentes STF. 6. Conduta que não pode ser tolerada, independentemente de quem a pratique, pois atenta contra o Poder Judiciário Eleitoral, colocando em risco a independência e a harmonia dos Poderes da República Federativa do Brasil e o próprio Estado Democrático de Direito. 7. Punição que deve ser, excepcionalmente, afastada no caso concreto em virtude do entendimento, pela maioria da Corte, quanto a fragilidade da prova nos autos de que a conduta do investigado foi em benefício de candidato, partido político ou coligação, no pleito eleitoral de 2018. 9. Ação de investigação judicial eleitoral julgada improcedente.

5. O teor do vídeo é inequívoco, residindo a controvérsia em questões de direito: legitimidade do pleito, possibilidade de enquadrar a conduta no art. 22 da LC 64/90 e gravidade dos fatos.
6. O sistema eletrônico de votação representa modelo de inegável sucesso implementado nas Eleições 1996 e internacionalmente reconhecido. O propósito dessa verdadeira revolução residiu na segurança e no sigilo do voto, sendo inúmeros os fatores que poderiam comprometer os pleitos realizados com urnas de lona, desde simples erros humanos na etapa de contagem, manipulações em benefício de candidatos e a execrável mercancia do sufrágio. Visou-se, ainda, conferir maior rapidez na apuração, o que possui especial relevância em países de dimensões continentais.
7. Esta Justiça Especializada não atua de forma sigilosa ou numa espécie de redoma na organização do pleito. Ao contrário, busca sempre soluções construtivas com os atores do processo eleitoral tendo como fim maior aperfeiçoar continuamente as eleições e consolidar o regime democrático.
8. A parceria entre órgãos institucionais de ponta na área de tecnologia, a constante busca por inovação e o contínuo diálogo com a sociedade propiciaram a plena segurança do sistema eletrônico de votação no decorrer dos últimos 25 anos, sem nenhuma prova de fraude de qualquer espécie, conforme inúmeras auditorias internas e externas e testes públicos de segurança diuturnamente noticiados pela Justiça Eleitoral.
9. Hipótese inédita submetida a esta Corte Superior é se ataques ao sistema eletrônico de votação e à democracia, disseminando fatos inverídicos e gerando incertezas acerca da lisura do pleito, em benefício de candidato, podem configurar abuso de poder político ou de autoridade – quando utilizada essa prerrogativa para tal propósito – e/ou uso indevido dos meios de comunicação quando redes sociais são usadas para esse fim.
10. Os arts. 1º, II e parágrafo único, e 14, § 9º, da CF/88, além dos arts. 19 e 22 da LC 64/90 revelam como bens jurídicos tutelados a paridade de armas e a lisura, a normalidade e a legitimidade das eleições. Não há margem para dúvida de que constitui ato abusivo, a atrair as sanções cabíveis, a promoção de ataques infundados ao sistema eletrônico de votação e à própria democracia, incutindo-se nos eleitores a falsa ideia de fraude em contexto no qual candidato sobrevenha como beneficiário dessa prática.
11. O abuso de poder político configura-se quando a normalidade e a legitimidade do pleito são comprometidas por atos de agentes públicos que, valendo-se de sua condição funcional, beneficiam candidaturas em manifesto desvio de finalidade. Precedentes.
12. Inviável afastar o abuso invocando-se a imunidade parlamentar como escudo. No caso de manifestações exteriores à Casa Legislativa a que pertence o parlamentar, "há necessidade de verificar se as declarações foram dadas no exercício, ou em razão do exercício, do mandato parlamentar; ou seja, se o denunciado expressou suas opiniões, sobre questões relacionadas a políticas governamentais; e se essas opiniões se ativeram aos parâmetros constitucionalmente aceitos, ou se teriam extrapolado eventuais parâmetros das imunidades materiais" (voto do Min. Alexandre de Moraes no Inquérito 4.694/DF, DJE de 1º/8/2019).
13. A internet e as redes sociais enquadram-se no conceito de "veículos ou meios de comunicação social" a que alude o art. 22 da LC 64/90. Além de o dispositivo conter tipo aberto, a Justiça Eleitoral não pode ignorar a realidade: é notório que as Eleições 2018 representaram novo marco na forma de realizar campanhas, com claras vantagens no uso da internet pelos atores do processo eleitoral, que podem se comunicar e angariar votos de forma mais econômica, com amplo alcance e de modo personalizado mediante interação direta com os eleitores.
14. No caso, constata-se sem nenhuma dificuldade que todas as declarações do recorrido durante sua live, envolvendo o sistema eletrônico de votação, são absolutamente inverídicas.
15. Quanto às urnas eletrônicas de seções eleitorais do Paraná, o recorrido atribuiu-lhes a pecha de "fraudadas", "adulteradas" e "apreendidas" e apontou que "eu tô com toda a documentação aqui da própria Justiça Eleitoral". Todavia, (a) inexistiu apreensão, mas mera substituição por problemas pontuais; (b) além da já enfatizada

segurança das urnas eletrônicas, a Corte de origem realizou auditoria antes do segundo turno – na presença de técnicos da legenda do candidato – e nada constatou; (c) é falsa a narrativa de que a suposta fraude estaria comprovada na "documentação aqui da própria Justiça Eleitoral", não havendo nenhuma menção a esse respeito nas atas das respectivas seções.

16. No tocante à declaração de que "nós não vamos aceitar que uma empresa da Venezuela, que a tecnologia que a gente não tem acesso, defina a democracia no Brasil", trata-se de inverdades refutadas inúmeras vezes: (a) sendo a Justiça Eleitoral criadora e desenvolvedora da urna eletrônica, seria no mínimo contraditório dizer que não há acesso à tecnologia de sistemas; (b) a empresa que produz as urnas não é venezuelana – o que, aliás, por si só, não representaria qualquer problema se fosse verdade.

17. É falsa a afirmativa de que apenas Brasil e Venezuela empregam urnas eletrônicas. Segundo o Instituto Internacional para a Democracia e a Assistência Social, 23 países as utilizam em eleições gerais e outros 18 em pleitos regionais, incluídos Canadá, França e algumas localidades nos Estados Unidos, o que também já foi esclarecido pela Justiça Eleitoral.

18. Inexistiu fraude nas Eleições 2014. Para além das inúmeras ocasiões em que a Justiça Eleitoral cumpriu com transparência seu dever de informação, houve auditoria externa conduzida pela grei derrotada naquele pleito, nada se identificando como irregular.

19. Os dividendos angariados pelo recorrido são incontroversos. A live ocorreu quando a votação ainda estava aberta no Paraná, ao passo que o acesso à internet ocorre de qualquer lugar por dispositivos móveis, reiterando-se que a transmissão foi assistida por mais de 70 mil pessoas, afora os compartilhamentos do vídeo.

20. O recorrido valeu-se das falsas denúncias para se promover como uma espécie de paladino da justiça, de modo a representar eleitores inadvertidamente ludibriados que nele encontraram uma voz para ecoar incertezas sobre algo que, em verdade, jamais aconteceu. Também houve autopromoção ao mencionar que era Deputado Federal e que a imunidade parlamentar lhe permitiria expor os hipotéticos fatos.

21. Gravidade configurada pela somatória de aspectos qualitativos e quantitativos (art. 22, XVI, da LC 64/90). O ataque ao sistema eletrônico de votação, noticiando-se fraudes que nunca ocorreram, tem repercussão nefasta na legitimidade do pleito, na estabilidade do Estado Democrático de Direito e na confiança dos eleitores nas urnas eletrônicas, utilizadas há 25 anos sem nenhuma prova de adulterações. Além disso, reitera-se a audiência de mais de 70 mil pessoas e, até 12/11/2018, mais de 400 mil compartilhamentos, 105 mil comentários e seis milhões de visualizações.

22. Na linha do parecer ministerial, "a transmissão ao vivo de conteúdo em rede social, no dia da eleição, contendo divulgação de notícia falsa e ofensiva por parlamentar federal, em prol de seu partido e de candidato, configura abuso de poder de autoridade e uso indevido de meio de comunicação", sendo grave a afronta à "legitimidade e normalidade do prélio eleitoral".

23. Recurso ordinário provido para cassar o diploma do recorrido e declará-lo inelegível (art. 22, XIV, da LC 64/90), com imediata execução do aresto, independentemente de publicação, e recálculo dos quocientes eleitoral e partidário.

(RECURSO ORDINÁRIO ELEITORAL nº 060397598, Acórdão, Relator(a) Min. Luis Felipe Salomão, Publicação: DJE - Diário de Justiça Eletrônico, Tomo 228, Data 10/12/2021)

Portanto, verifica-se o aludido julgamento representando um marco, por se tratar da primeira cassação de um candidato por disseminar *fake news*. Destaca-se a Corte Eleitoral manifestando forte valorização do sistema de votação da Justiça Eleitoral brasileira, além de enaltecer a solidez do processo eleitoral. Por outro lado, o julgado representa uma reafirmação da jurisprudência do TSE, quanto a interpretação de considerar redes sociais como meio de

comunicação social, passível de configurar ilícito de uso indevido, conforme prevê o Art. 22, da LC n. 64/1990.

Destarte, com a disseminação de desinformação (*fake news*), passa-se a considerar, além das questões pertinentes à propaganda eleitoral, como os pedidos de remoção de conteúdo e pedidos de direitos de resposta, também, a possibilidade de ações passíveis de culminar com a cassação de candidatos, considerando o TSE, agora, entendendo as redes sociais como meios de comunicação social, sua utilização para propagar desinformação poderá resultar em cassação de registro, diploma ou mandato, bem como ser considerado inelegível, pelo prazo legal de 8 (oito) anos, iniciando a contagem do prazo na data da eleição que se constatou o ilícito.⁶²²

⁶²² ALARCON, Anderson de Oliveira. **Sociedade da (des)informação**: análise da jurisprudência e das decisões do TSE sobre fake news nas eleições de 2018 e de 2020. 2022. Tese de Doutorado – Universidade Federal do Paraná, Setor de Ciências Humanas, Programa de Pós-Graduação em Ciência Política. Curitiba, 2022. p. 115.

CONSIDERAÇÕES FINAIS

Extraí-se da pesquisa os seguintes apontamentos finais, organizados em tópicos objetivos e sintéticos, sem prejuízo de outras conclusões e considerações já esboçadas ao longo do trabalho.

1. No Capítulo I, abordou-se a liberdade de expressão, seja no ambiente *online*, ou no *offline*, a partir do entendimento de um conjunto de direitos e faculdades individuais de exprimir, de informar, ser informado, de não se expressar ou de não se informar. Assumindo, deste modo, uma dualidade entre a liberdade de expressão estrito senso, não ligada à veracidade da opinião, e, por outro lado, a liberdade de informação limitada pela verdade e pela realidade fática. Constatou-se, ainda, que no Brasil é previsto, constitucionalmente, um direito geral de liberdade, o qual reconhece, além das liberdades específicas, as de expressão, reunião, opinião, e manifestação, servindo como objeto de critério material para reconhecimento de outras liberdades.

2. Nesse sentido, a liberdade de expressão é considerada imprescindível e constitui um direito fundamental precioso, figurando nos catálogos de direitos fundamentais desde o advento do constitucionalismo moderno. Por se tratar de um importante instrumento para a descoberta da verdade, torna-se imprescindível à participação democrática.

3. Foram abordados os valores necessários à manutenção de um sistema de proteção da liberdade de expressão, assegurando a satisfação individual, além de uma forma de alcançar a verdade, buscar um método para garantir a participação da sociedade no debate social e político, servindo como uma forma de equilibrar a estabilidade e a mudança da sociedade.

4. Examinadas as teorias argumentativas da liberdade de expressão como direito fundamental, chegou-se ao entendimento de a liberdade de expressão possuir natureza multifuncional, constituindo um sistema. Assim, os argumentos instrumentais e constitutivos não se excluem mutuamente, tal como defendido por Jónatas Machado.

5. Quanto à concepção de direitos fundamentais limitados ou restringíveis, entendemos o predomínio e entendimento de não serem absolutos, decorrendo, principalmente, do caráter universal ou geral dos direitos fundamentais, bem como da necessidade de esses direitos coexistirem entre si, ou com outros direitos fundamentais protegidos constitucionalmente. Sendo, assim, foi demonstrado o debate entre a limitação da liberdade de expressão, optando entre interna ou externa, nesse ponto, adotamos no trabalho a teoria da liberdade de expressão externa.

6. Tomando-se, pois, que, na teoria externa, consideram-se os direitos fundamentais limitados pelos próprios direitos fundamentais de outra natureza. Assim, a liberdade de expressão tem como limite externo os direitos da personalidade, tais como: a verdade, a honra e a privacidade, representando alguns desses limites.

7. Constatou-se que a liberdade de expressão *offline* passa a ser também *online*, a qual chamamos de digitalização da fala ou do discurso, representando uma elevação de sua proeminência no constitucionalismo moderno. Esse processo de digitalização trouxe vários benefícios para governos, cidadãos e empresas nas últimas décadas. Contudo, essa digitalização é um conceito abrangente, compreendendo a relação ente cidadãos e o governo e destes com empresas privadas. Nesse ponto, o processo foi integrado por plataformas digitais, as quais oferecem diferentes serviços e produtos.

8. Verificamos que as plataformas digitais desempenham um papel importante e fundamental na economia atual. Entretanto, observam-se algumas plataformas digitais não se ocupando apenas com seu modelo de negócio e com o lucro dali advindo, havendo expansão do seu poder, cada dia mais presente em vários setores da sociedade. Com isso, apontamos no estudo que as plataformas buscam, além de seu retorno econômico, estabelecer uma agenda política, fazendo com que as novas questões jurídicas surjam, pois, como previsto na relação triangular de Balkin, é necessária uma nova forma de governança de discurso *online*. Finalizando o Capítulo I, atingimos o objetivo de *fornecer* um panorama da liberdade de expressão no ambiente virtual e suas principais implicações.

9. A partir desse olhar, o estudo analisou a governança da *Internet* no Capítulo II, inserido num debate da regulação de serviços prestados através da *Internet*, possuindo contornos de embate público em relação às grandes empresas privadas as quais exploram a infraestrutura *online*. O trabalho não se ocupou de tratar das formas de regulação, nem em apontar qual deveria ser o tipo adotada. O exame concentrou-se na estrutura da governança, apontando os integrantes desse meio ambiente, especialmente, os atores privados, dentre eles as empresas de redes sociais.

10. Concentramos-nos na camada de conteúdo, onde se situam os símbolos, imagens, áudios e demais multimídias, dentre tais camadas da *Internet* existem – a física, de código e de conteúdo, bem como os dispositivos que interagem diretamente com os conteúdos os quais trafegam. Nessa camada, em cujos principais resultados da escolha do *design* são percebidos, implicando as questões de liberdade de expressão, privacidade, censura privada, acesso ao conhecimento, além de outros.

11. Entendemos a Governança da *Internet* como uma estrutura que reúne diversos atores, processos, tecnologias e infraestrutura, interagindo entre si, formando uma complexa e plural atividades no mundo digital. Em resumo, é um conjunto dinâmico de disputas e deliberações sobre como a *Internet* é coordenada, administrada e moldada para refletir políticas públicas (Mueller).

12. Para melhor compreender a estrutura da governança da *Internet*, adotamos a abordagem em duas dimensões: subjetiva e objetiva (Keller). Na dimensão subjetiva, compreende-se o grupo de atores, integrando os processos e a gestão os quais influenciam na formação de políticas no ambiente *online*. Inicialmente, era composta por instituições intergovernamentais e multissetoriais administrando e coordenando os recursos da *Internet*, dentre eles abordamos o ICANN, IETF, IANA, W3C. Atualmente, esse grupo integra também os fóruns de debates, além de organizações com influência no comportamento *online*, como organizações da sociedade civil, Estados-nação e as empresas privadas.

13. Enquanto a dimensão objetiva, representa a forma de como a governança da *Internet* e se materializa na prática, por meio de várias práticas, dentre elas a coordenação de nome de domínios realizadas pelas pelo ICANN; na elaboração de leis nacionais, objetivando a neutralidade de rede, dispondo sobre a não discriminação de conteúdo por provedores de acesso; pela criação de filtro de conteúdo impróprios para crianças, por provedores de serviços; além da exclusão de conteúdo ou contas de usuários, consideradas impróprios por determinar rede social; na camada de conteúdo, questões ligadas à interoperabilidade e interface, relacionando-se com a privacidade, liberdade de expressão, direitos de propriedade intelectual.

14. Após a análise da governança da *Internet*, quanto à sua estrutura e composição, passamos a destacar a importância com a qual os atores privados, em especial as empresas exploradoras da infraestrutura *online*, tem ocupado nesse cenário. Exercendo uma governança privada em suas plataformas, sendo então denominada de “governança das plataformas” (Gorwa). Nesse ponto, destacamos a necessidade de cotejar os termos de serviços de forma geral, tem-se construído, com algum aspecto constitucional em suas normas, e verificando se há observância dos direitos fundamentais em sua constituição.

15. Constatou-se ser inegável que as redes sociais representaram um avanço, inclusive, democrático, na sociedade da informação. Elevando as plataformas a um *status* próximo de Estado, ditando suas normas e condições, rumo a seus desideratos os quais, nem sempre estão associados ao bem-comum. Surgindo, inclusive, o debate que ocorrem em Estados tradicionais, nos quais encontram um desafio democrático, com a evolução dos atores

privados, determinando, de certa forma, os serviços e as políticas públicas, exercendo influência direta nos usuários e criando suas regras, as quais nem sempre são iguais às tradicionais do Estado.

16. Nessa linha, foi examinada a tendência de os termos de uso assumirem um tom constitucional, em especial, a rede social *Facebook*. De fato, entendemos que estes termos transcendem de um mero contrato privado entre as partes, apesar de não possuir uma força vinculante, tem assumido um papel de regramento da política da comunidade, atingindo direitos fundamentais dos usuários, sobretudo, na tomada de decisões quanto à exclusão de conteúdos e usuários da plataforma. Contudo, encampamos o posicionamento do não reconhecimento da capacidade de as entidades privadas desempenharem as duas funções constitucionais supremas, sendo elas, a proteção dos direitos fundamentais e o equilíbrio dos poderes desses atores envolvidos (Celeste). Com efeito, concordamos com os documentos constituintes de um processo de constitucionalização do ambiente das mídias sociais. Objeto a ser estudado e acompanhado pela academia, tal como já é desenvolvido pelo estudo do Constitucionalismo Digital.

17. Concluímos a partir desse ponto, a necessidade da democratização dos termos de serviço das plataformas *online*. Pois, como demonstrado, os termos de uso ou de serviço incluem cláusulas as quais transcendem a exigência básica para a moderação do conteúdo pelas plataformas, com isso, é de suma importância a participação do usuário na formulação desses termos. O estudo sugere que essa construção conjunta entre plataforma e usuários estabelecerá termos de uso com maior respaldo e legitimidade, perante os membros da comunidade, por ser uma construção de consenso.

18. As implicações com os direitos humanos, especialmente a liberdade de expressão, encontradas no ambiente *online* também foram objeto de estudo do Capítulo II onde a primazia do controle pertence a atores privados. Seria dizer que se ocupou do conflito no plano *horizontal*, entre o intermediário, impondo a medida e o usuário final. Apontamos, como meio alternativo para alcançar a responsabilidade dos direitos humanos pelos atores sociais, a observância dos “Princípios Orientadores sobre Negócios e Direitos Humanos das Nações Unidas”, bem como analisamos o balanço de 10 anos dos Princípios Orientadores sobre Negócios e Direitos Humanos. Constata-se, em síntese, a persistência de abusos relacionados às empresas, devendo assumir atenção prioritária urgente por parte dos Estados e das empresas. Verificou-se o destaque, que, na última década, ressaltou o ponto levantado nos Princípios Orientadores: abordagens voluntárias, por si só não são suficientes. É necessário o aumento das medidas obrigatórias para acelerar a aceitação e o progresso. Bem como, as

medidas legais são essenciais, mas não suficientes para garantir o respeito das empresas pelos direitos humanos.

19. Nessa perspectiva, apresentamos o resultado da pesquisa realizada pelo CTS/FGV – Centro de Tecnologia e Sociedade da Escola de Direito do Rio de Janeiro da Fundação Getúlio Vargas, em parceria com a *Coalizão Dinâmica sobre a Responsabilidade das Plataformas* (DCPR), realizando a análise da adequação dos contratos das plataformas *online*, em relação aos direitos humanos, baseado na estrutura do Guia de Direitos Humanos para os Usuários da *Internet*.

20. Outro ponto tratado, foi o Primeiro Relatório de Direitos Humanos da *Meta*, empresa detentora de várias plataformas, dentre elas, o *Facebook*. O relatório é baseado no Princípio 15, dos Princípios Orientadores da ONU sobre Empresas e Direitos Humanos. Apontamos que o Relatório deixa de tratar de alguns temas, outros são abordados insuficientes. Entretanto, entendemos que representa um avanço, pois a *Meta* reconheceu a necessidade de comprometimento com uma política de direitos humanos. Ademais, constatou-se que o relatório é lacônico quanto às políticas de moderação de conteúdo. Com isso, fica lançada a pretensão de a empresa aderir aos princípios de direitos humanos. É importante o reconhecimento de o primeiro passo ter sido dado ao reconhecer a existência do problema.

21. Finalizando o capítulo, dentro da teoria dos direitos fundamentais, tratamos da eficácia nas relações entre particulares, enfatizada na associação entre as plataformas *online* e os seus usuários, nós nos apoiamos na concepção majoritária de reconhecimento à eficácia direta *prima facie* dos direitos fundamentais, especialmente o da liberdade de expressão. Observando a existência de entendimentos contrários, nos quais são adotadas a eficácia indireta. A Constituição Federal de 1998 dispõe, expressamente, no Art. 5º, § 1º, que as normas definidoras dos direitos e garantias fundamentais possuem aplicação imediata. Que a nosso ver – bem como a doutrina majoritária – demonstra ser uma decisão patente em reconhecer a eficácia direta das normas de direitos fundamentais.

22. O estudo encampa a proposta de um novo pensamento sobre o tipo de eficácia mais adequada a ser aplicada nas relações entre particulares. Consistente numa aplicação simultânea da eficácia indireta, como regra e da eficácia direta de casos singulares (Sarlet e Hartmann). A eficácia indireta, adotando a teoria dos deveres de proteção, passando o legislador a produzir normas específicas protetora dos direitos fundamentais dos autorreguladores e daqueles sujeitos aos efeitos dessa autorregulação. E em casos específicos, aplicar a eficácia direta em relações como redes sociais privadas, ocupantes de lugar privilegiado na relação com usuários, possibilitando a realização de condutas violadoras dos

direitos fundamentais como a liberdade de expressão, além de outros. Neste Capítulo II, alcançamos o objetivo de *compreender* como a governança privada das plataformas digitais, analisando atores privados, sobretudo quanto às regras elaboradas pelas plataformas *online* e os direitos fundamentais, especialmente a liberdade de expressão.

23. No Capítulo III, foram abordados os principais tópicos relacionados ao debate de moderação de conteúdo, para contextualizar a discussão da matéria, fazendo o recorte de remédios alternativos a serem estudados quanto à remoção de conteúdo, cuja aplicação se dá após a violação. Portanto, depois da publicação do conteúdo, tem-se a moderação *ex post*, a qual é centrada nas ações dos atores privados, não sendo objeto deste trabalho as ações de atores estatais ou do sistema jurídico.

24. Compreendemos a moderação como os mecanismos de governança que estruturam a participação em uma comunidade para facilitar a cooperação e prevenir abusos (Grimmelmann). Percebeu-se, deste modo, a moderação de conteúdo como fator essencial à operação das plataformas de redes sociais. Sendo assim, foi abordado o papel do moderador das empresas de mídia social, tanto pela implementação de diretrizes adotadas, quanto pelo processo ativo de moderação de conteúdo gerenciados em relação a reclamações específicas, bem como o cumprimento das leis sobre conteúdo ilegal, como violação das regras internas.

25. Quanto à aplicação dos remédios na moderação de conteúdo, defendemos uma maior diversificação. Entre as soluções existentes, além da remoção, codificada, inicialmente, pelos reguladores como a solução primordial, por se tratar de solução primária ou exclusiva pelas plataformas *online*. Posteriormente, ao entendê-la como solução secundária, subsidiária a um regramento mais completo e encorpado, verificou-se a necessidade de soluções mais definitivas, não aquelas cuja função não seria a cura, apenas um “band-aid” para um problema já existente.

26. Com efeito, reconhecemos a remoção de conteúdo como parte do *kit* de ferramentas de correção dos serviços de *Internet*. Contudo, nossa defesa é que não seja o único mecanismo, tampouco o mais importante. Nesse contexto, havendo e se reconhecendo instrumentos os quais sirvam de remédios alternativos à remoção. É necessário trazer em relevo os possíveis danos causados por conteúdos transgressores, a fim de balanceá-los com os benefícios da permanência de determinada postagem. Exatamente por isso, ao sopesar tais valores, percebe-se a relevância e a promoção da liberdade de expressão, esse leque de ferramentas ainda corrige o conteúdo que viola as regras da plataforma.

27. Com essa visão, restou demonstrado a dificuldade de moderação de conteúdo. Sendo praticamente impossível a tarefa de atenuar o teor de postagens, por meio da

moderação, de forma a atender a todos os interesses, satisfazendo a plataforma, aos usuários e o Estado, enfim, a todos. Desta forma, o estudo propõe uma diversidade de soluções de moderação de conteúdo, com opções interessantes e, até mesmo, simples, auxiliando os serviços de *Internet* para moderar suas comunidades e ponderar os vários interesses concorrentes.

28. O trabalho adotou a taxonomia de opções de remédios de Eric Goldman, divididas em cinco categorias: 1. ações em face de parte de conteúdo individuais; 2. ações contra conta *online*; 3. ações para reduzir a visibilidade das violações, tanto para conteúdos individuais ou conta inteira; 4. ações impondo consequências financeiras por violações, tanto em face de conteúdos individuais ou conta inteira; e 5. outras ações as quais não se enquadrem nas categorias anteriores.

29. Apresentamos os principais remédios, visando corrigir violações, evitando – ao máximo – a remoção definitiva de conteúdo ou de contas. Consignando como os remédios podem serem aplicados tanto de forma autônomas, como também de forma combinada, sempre prestigiando a manutenção do conteúdo/conta. Evidenciamos como nenhum remédio será eficaz para a maioria das correções, pois são inúmeras as formas e, a cada dia, surgem novas formas de violação. Além disso, muitos remédios são eficazes para um determinado tipo de plataformas *online*, enquanto para outras são insuficientes.

30. Com a demonstração dessas soluções alternativas, buscamos contribuir para a qualidade dos processos de moderação de conteúdo, alcançando modelos de moderação de conteúdo centrados nos direitos humanos reconhecendo os direitos dos usuários.

31. Alinhado a esses recursos alternativos, entendemos o auxílio prestado por eles como medida de transparência, devendo integrar qualquer sistema de moderação de conteúdo. Estabelecendo, com isso, uma confiança entre os usuários e os moderadores de conteúdo. Sugerimos, ainda, como forma de aumentar a transparência das moderações, a publicação de relatórios de transparência pelas plataformas de acordo com cada remédio alternativo empregado. Assim, ficaria mais claro se a moderação do conteúdo foi realizada com base na exigência legal, ou se foi baseada nas normas internas da plataforma *online*, ou seja, os termos de uso. Além de tornar explícita as decisões tomadas, entre remover ou adotar algum remédio alternativo de moderação a manutenção do conteúdo na plataforma.

32. A moderação de conteúdo no Brasil, realizado pelas grandes plataformas de mídia social, assume sobre aspectos distintos e próprios, sobretudo, em decorrência da previsão normativa regida pelo Marco Civil da Internet (Lei n. 12.965/2014). O diploma legal não dispõe exclusivamente sobre a moderação de conteúdos por plataformas de redes sociais,

englobando outros intermediários. O Artigo 19 se destaca ao dispor no *caput* a busca para assegurar a liberdade de expressão e impedir a censura, estabelecendo a responsabilidade civil dos provedores de aplicações, em decorrência de conteúdos criados por terceiros, somente incidirá após ordem judicial específica, de remoção de conteúdo.

33. Identificamos que o Marco Civil da Internet trata apenas da governança estatal, silenciando quanto à governança privada pelos intermediários. Propondo que a remoção de conteúdo é prerrogativa exclusiva do poder judiciário. Nesse aspecto, apresentamos duas conclusões: na primeira, não vislumbramos nenhum óbice quanto à prerrogativa “tradicional” do Poder Judiciário em dirimir conflitos em relação à legalidade ou não de um determinado conteúdo. Já na segunda, apontamos como temerário não prever a moderação de conteúdo realizada pelas próprias plataformas *online*. Deixando esse árduo processo ao poder judiciário, hoje, bastante sobrecarregado e burocrático, causando ainda mais lentidão na solução das demandas, não sendo condizente com a fluidez e a velocidade da produção de conteúdo em ambientes *online* e o quão ágil os moderadores devem ser para manter o ambiente digital saudável, aberto e democrático.

34. Destarte, entendemos a necessidade de haver, de um lado, os mecanismos tradicionais do direito do Estado, materializando-se no Marco Civil da Internet no Brasil e, por outro lado, os remédios alternativos de moderação de conteúdo, aplicados por empresas de mídias sociais, através dos arranjos de governança privada de moderação de conteúdo. Cabendo ao ordenamento jurídico brasileiro, reconhecer a autonomia para tomada de decisões pelas plataformas *online*, com a incidência de direitos fundamentais nas relações entre particulares, especialmente, a liberdade das empresas em definirem regras de discursos em seus ambientes, observando a liberdade de expressão dos usuários. Embora possibilite as plataformas tomarem suas decisões no bojo da moderação de conteúdo, elas estarão sujeitas à análise Judiciário e poderá decidir quanto à licitude ou constitucionalidade, bem como examinar a razoabilidade da medida de restrição, nos casos envolvendo a eficácia horizontal de direitos fundamentais, no qual haja pressuposição de uma dimensão de autonomia entre as partes privadas. Nesse Capítulo III, realizamos o objetivo de *identificar* os principais tópicos de discussão sobre a moderação de conteúdo pelas plataformas *online*, apresentar rol de remédios para correção de conteúdo ou ações transgressores das regras internas das empresas de mídia social, para além da simples remoção de conteúdo.

35. Logo, em resposta ao problema da pesquisa, o trabalho evidenciou a possibilidade de as plataformas de redes sociais removerem diretamente os conteúdos e/ou contas de seus usuários, nos casos de eles violarem suas normas internas, a fim de manter seu ambiente

saudável, aberto e democrático. Contudo, a exclusão não é único meio. Em nossa análise observou-se ainda a existência de outros mecanismos de moderação, devendo tais remédios serem utilizados, prioritariamente, à remoção do conteúdo ou outro remédio alternativo para corrigir tal violação.

Assim, restou demonstrado a existência de outras alternativas para solucionar os problemas de moderação de conteúdo, para além da simples remoção. Assim, como forma de proteger a liberdade de expressão dos usuários, as plataformas precisam protegê-los, além de fazê-los sentirem-se protegidos no ambiente preparado para eles. Esse espaço *online* deve ser livre, confiável, saudável, prazeroso e um lugar em que os usuários possam se expressar sem ser invadido, mas sem invadir o espaço de outrem, propício, dessa forma, ao debate *online*. Se esse espaço democrático permitir o exercício da liberdade de expressão a todos, independentemente, de religião, preferências sexuais, nacionalidade, cor da pele ou gênero.

36. Desta forma, concluímos o trabalho apontando a governança da fala *online* como um processo extremamente difícil e complexo, que talvez nunca atinja a um ponto de chegada ideal, pois, a moderação de conteúdo deverá estar sempre em fase de melhoria contínua, atendendo aos critérios internos da empresa, bem como do ambiente de negócio. Não há, portanto, moderação de conteúdo pronta e acabada, com normas estáveis ou com processos regulatórios definitivos.

Percebe-se então que a as plataformas depende de um sistema de cooperação mútuo em que os atores particulares (os proprietários do ambiente virtual e o usuário) e públicos contribuam para esse padrão de constante melhoria para buscar, movimento e melhoria contínua, respeito às regras do negócio, adaptações e experimentações, sempre movido pelo desenvolvimento tecnológico. Todo esse processo de gigantes no controle de muitas vozes, sobretudo após o período pandêmico, mostra que vivenciamos uma época disruptiva e inédita, transparecendo que os meios de governança da *Internet* são contestadas e se tornaram ineficazes, e novos arranjos estão surgindo para supri-las.

37. Com isso, a presente pesquisa representa mais uma contribuição para o avanço da compreensão dessas novas tarefas obrigatórias das empresas de mídia social. Projetando, para tanto, dois pontos importantes para o futuro da moderação de conteúdo, tornando-se objeto de novos estudos: primeiro, a necessidade de uma *recalibragem da moderação de conteúdo*, sobretudo, com a nova realidade mostrada pela pandemia da COVID-19, devendo ser adotado um *equilíbrio sistêmico* na governança do discurso *online*, deixando-se de lado a atuação *individualizada* da moderação de conteúdo. Em segundo lugar, a premente necessidade de um exame da *moderação de conteúdo* atual seguindo um sistema *formalista*, devendo evoluir

para adotar um sistema *realista*. Além de outros pontos, cuja bibliografia deste trabalho pode ser útil.

REFERÊNCIAS

ABBATE, Janet. **Inventing the Internet**. Cambridge: MIT Press, 1999.

ABBOUD, Georges; NERY JÚNIOR, Nelson; CAMPOS, Ricardo. **Fake news e regulação**. São Paulo: Thomson Reuters Brasil, 2021.

ALARCON, ANDERSON DE OLIVEIRA. **Sociedade da (des)informação**: análise da jurisprudência e das decisões do TSE sobre fake news nas eleições de 2018 e de 2020. 2022. Tese de Doutorado – Universidade Federal do Paraná, Setor de Ciências Humanas, Programa de Pós-Graduação em Ciência Política. Curitiba, 2022.

ALEXANDER, Larry. **Is There a Right of Freedom of Expression?** New York: Cambridge University Press, 2005.

ALEXY, Robert. **Teoria dos Direitos Fundamentais**. Tradução Vírgilio Afonso da Silva. 2ª ed. São Paulo: Malheiros Editores, 2011.

ANDREWS, Leighton. Facebook ‘Regulation’: a process not a text. **UK Copyright Centre and Creative Economy Centre (CREATE)**. 2020. Disponível em: <https://orca.cardiff.ac.uk/id/eprint/133226/1/CREATE%20Text.pdf>. Acesso em: 03/11/2021.

ARROYO, César Landa. Derecho fundamental al Internet: contenido esencial. **Pensar**. Fortaleza. v. 23, n. 4, p. 1-22, 2018. Disponível em: <https://periodicos.unifor.br/rpen/article/view/8447/pdf>. Acesso em: 24/10/2021.

ARUN, Chinmayi. Facebook’s Faces. **Harvard Law Review Forum**. v. 135, p. 236-264, 2022. Disponível em: <https://harvardlawreview.org/wp-content/uploads/2022/03/135-Harv.-L.-Rev.-F.-236.pdf>. Acesso em: 30/06/2022.

ASH, Timothy Garton. **Free Speech: Ten Principles for a Connected World**. London: Atlantic Book, 2016.

BACA, Vilmar Villena Fernández. Derecho fundamental de acceso a internet: Um acercamiento preliminar a sus contenidos. **Gaceta Constitucional**. N. 163, p. 194-204, 2021. Disponível em: https://www.academia.edu/50343496/Derecho_fundamental_de_acceso_a_internet_Un_acercamiento_preliminar_a_sus_contenidos. Acesso em 26/08/2021.

BAKER, C. Edwin. **Human Liberty and Freedom of Speech**. New York: Oxford University Press, 1989.

BALKIN, Jack M. Digital Speech and Democratic Culture: A Theory of Freedom of Expression for the Information Society, **New York University Law Review**, v. 79, p. 1–55, 2004. Disponível em: https://openyls.law.yale.edu/bitstream/handle/20.500.13051/1712/Digital_Speech_and_Democratic_Culture_A_Theory_of_Freedom_of_Expression.pdf?sequence=2. Acesso em 04/11/2021.

BALKIN, Jack M. Free Speech is a triangle. **Columbia Law Review**. v. 118, p. 2011–2055, 2018. Disponível em: <https://openyls.law.yale.edu/bitstream/handle/20.500.13051/17953/Balkin%2C%20Free%20Speech%20is%20a%20Triangle.pdf?sequence=1&isAllowed=y>. Acesso em: 04/11/2021.

BALKIN, Jack M. Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation. **University of California Davis Law Review**. Volume 51, n. 3, p. 1149–1210, 2018. Disponível em: https://lawreview.law.ucdavis.edu/issues/51/3/Essays/51-3_Balkin.pdf. Acesso em: 04/11/2021.

BALKIN, Jack M., How to Regulate (and Not Regulate) Social Media. **Journal of Free Speech Law**. v. 1, p.71-96, Knight Institute Occasional Paper Series, n. 1, Yale Law School, Public Law Research Paper. 2021. Disponível em: <https://ssrn.com/abstract=3484114>. Acesso em: 10/12/2021.

BALKIN, Jack M. Old-School/New-School Speech Regulation. **Harvard Law Review**. v. 127, n. 8, p. 2296-2342., 2014. Disponível em: https://openyls.law.yale.edu/bitstream/handle/20.500.13051/4406/127_Harv_L_Rev._2296_2014.pdf?sequence=2&isAllowed=y. Acesso em: 04/11/2021.

BALKIN, Jack M. The Future of Free Expression in a Digital Age, **Pepperdine Law Review**, v. 36, Issue 2, 2009. Disponível em: <https://digitalcommons.pepperdine.edu/plr/vol36/iss2/9/>. Acesso em: 04/11/2021.

BARCELOS, Júlia Rocha de. **Big Data, Algoritmos e Microdirecionamento**: desafios para a regulação da propaganda eleitoral. Dissertação (Mestrado). Universidade Federal de Minas Gerais, Faculdade de Direito, Belo Horizonte: 2019;

BARENDT, Eric. **Freedom of speech**. 2 ed. New York: Oxford University Press, 2ª ed. 2005.

BARLOW, John P. **Uma Declaração de Independência do Ciberespaço**, Electronic Frontier Found. (8 de fevereiro de 1996), Disponível: www.eff.org/cyberspace-independence. Acesso em: 29/10/2021.

BARROSO, Luna van Brussel. **Liberdade de expressão e democracia na Era Digital**: o impacto das mídias sociais no mundo contemporâneo. Belo Horizonte: Fórum, 2022.

BELLI, Luca. Governança e regulações da internet: uma apresentação crítica. In: BELLI, Luca; CAVALLI, Olga (Org.). **Governança e regulações da Internet na América Latina**: análise sobre infraestrutura, privacidade, cibersegurança e evoluções tecnológicas em homenagem aos dez anos da South School on Internet Governance. Rio de Janeiro: Escola de Direito do Rio de Janeiro da Fundação Getúlio Vargas, 2019.

BELLI, Luca; FRANCISCO, Pedro Augusto; ZINGALES, Nicolo. Law of the Land or Law of the Platform? Beware of the Privatisation of Regulation and Police. p. 41-64. In: BELLI, Luca; ZINGALES, Nicolo. **Platform regulations**: how platforms are regulated and how they regulate us. Official outcome of the UN IGF Dynamic Coalition on Platform Responsibility. Rio de Janeiro: Escola de Direito do Rio de Janeiro da Fundação Getúlio Vargas, 2017.

BELLO, Pablo; SASTRE, Andrés. Repensar as políticas públicas para fechar o fosso digital na América Latina. In: BELLI, Luca; CAVALLI, Olga (Org.). **Governança e regulações da Internet na América Latina**: análise sobre infraestrutura, privacidade, cibersegurança e evoluções tecnológicas em homenagem aos dez anos da South School on Internet Governance. Rio de Janeiro: Escola de Direito do Rio de Janeiro da Fundação Getúlio Vargas, 2019. p. 247-266.

BIANCHI, José Flávio. Internet e Regulação: a ICANN à luz da teoria da regulação. **Revista de Direito, Estado e Telecomunicações**. Brasília, v. 8, n. 1, p. 135-156, maio 2016. Disponível em: <https://doi.org/10.26512/lstr.v8i1.21529>, Acesso em: 16/06/2022.

BINENBOJM, Gustavo. **Liberdade igual**: o que é e por que importa. 1 ed. Rio de Janeiro: História Real: 2020.

BORK, Robert H. Neutral Principles and Some First Amendment Problems. **Indiana Law Journal**, Vol. 47, Fall 1971, Number I. Disponível em: <https://heinonline.org/HOL/LandingPage?handle=hein.journals/indana47&div=8&id=&page>. Acesso em: 28/12/2021.

BOROWSKI, Martin. La restricción de los derechos fundamentales. **Revista Española de Derecho Constitucional**. Ano 20, n. 59, p. 29-56, 2000. Disponível em: <https://dialnet.unirioja.es/servlet/articulo?codigo=79684>. Acesso em: 10/01/2022.

BOWERS, Jonh; ZITTRAIN, Jonathan. Answering impossible questions: content governance in an age of disinformation. **Harvard Kennedy School (HKS) Misinformation Review**. v. 1, n. 1. 2020. Disponível em: <https://misinforeview.hks.harvard.edu/article/content-governance-in-an-age-of-disinformation/>. Acesso em: 08/08/2022.

BRADSHAW, Samantha; DENARDIS, Laura; HAMPSON, Fen; JARDINE, Eric; RAYMOND, Mark. **The Emergence of Contention in Global Internet Governance**. 56ª Convenção Anual da ISA. New Orleans, USA. 2015. Disponível em: <http://web.isanet.org/Web/Conferences/New%20Orleans%202015/Archive/adf6039d-3251-4ede-a2fa-0f8701f21459.pdf>. Acesso em: 25/06/2022.

BRANCO, Paulo Gustavo Gonet. Liberdades. In: MENDES, Gilmar Ferreira; BRANCO, Paulo Gustavo. **Curso de direito constitucional**. 15. ed., São Paulo: Saraiva, 2020.

BRANCO, Paulo Gustavo Gonet; BRANCO, Pedro Henrique de Moura Gonet. Structural Way of Governing the Internet – The Brasilia Position. **Revista Jurídica**. v. 2, n. 55, p. 592-606, Curitiba, 2019. Disponível em: <http://revista.unicuritiba.edu.br/index.php/RevJur/article/view/3840/371372180>. Acesso em: 06/11/2021.

BRASIL. Decreto nº 4.829, de 3 de setembro de 2003. Dispõe sobre a criação do Comitê Gestor da Internet no Brasil - CGIbr, sobre o modelo de governança da Internet no Brasil, e dá outras providências. **Diário Oficial da República Federativa do Brasil**, Poder Executivo. Brasília, 2003.

BRUZZONE, Andrés. **Ciberpopulismo: política e democracia no munda digital**. São Paulo: Contexto, 2021.

CALLAMARD, Agnès. The Human Rights Obligations of Non-State Actors. In: JØRGENSEN, Rikk Frank [Ed.]. **Human Rights in the Age of Platforms**. Cambridge, MA: The MIT Press, 2019. Disponível em: <https://direct.mit.edu/books/book/4531/Human-Rights-in-the-Age-of-Platforms>. Acesso em: 29/10/2021

CALLEJÓN, Francisco Balaguer. Redes Sociais, Companhias Tecnológicas e Democracia. **REI - REVISTA ESTUDOS INSTITUCIONAIS**, [S.l.], v. 6, n. 2, p. 579-599, set. 2020. Disponível em: <https://www.estudosinstitucionais.com/REI/article/view/485> ou <https://doi.org/10.21783/rei.v6i2.485>. Acesso em: 26/06/2022.

CAMPO, Agustina Del; SCHATZKY, Morena; HERNÁNDEZ, Laura. LARA, J. Carlos. **Olhando Al Sur**. Rumo a novos consensos regionais em matéria de responsabilidade de intermediários na Internet. 2021. Disponível em: <https://www.alsur.lat/sites/default/files/2021-06/Responsabilidad%20de%20intermediarios%20PT.pdf>. Acesso em: 02/08/2022.

CANABARRO, Diego. Governança Global da Internet: Aspectos Conceituais, Questões da Agenda Contemporânea e Prospectos para o Estudo do Tema. In: OPPERMANN, Daniel. (Ed.). **Internet Governance in the Global South: History, Theory, and Contemporary Debates**. São Paulo: University of São Paulo, 2018. Disponível em: <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-65805-2>. Acesso em: 08/07/2022.

CANARIS, Claus-Wilhelm. Considerações a respeito da posição de proibições de discriminação no sistema do direito privado. **Revista Brasileira De Direitos Fundamentais & Justiça**, v. 7, n. 22, p. 15–20, 2013. Disponível em: <https://doi.org/10.30899/dfj.v7i22.279>. Acesso em: 10/12/2021.

CANOTILLO, J. J. GOMES. **Direito constitucional e teoria da constituição**. 7ª ed., 21 reimp. Coimbra: Almedina, 2019.

CANOTTILHO, José Joaquim Gomes; MACHADO, Jónatas E. M; GAIO JÚNIOR, Antônio Pereira. **Biografia não autorizada versus liberdade de expressão**. 3. ed. Curitiba: Juruá, 2017.

CANOTILHO, José Joaquim Gomes; MACHADO, Jónatas E. M. **“Reality Shows” e a Liberdade de Programação**. Coimbra: Coimbra Editora, 2003.

CAROTTI, Bruno. ICANN and Global Administrative Law. In: CASSESE, Sabino et alli. (eds.). **Global Administrative Law: Cases, Materials, Issues**. New York University School of Law: Institute for International Law and Justice, 2008. Disponível em: <https://studylib.net/doc/15758957/icann-and-global-administrative-law--by-bruno-carotti>. Acesso em: 29/05/2022.

CARVALHO, Luis Gustavo Grandinetti Castanho de. **Liberdade de informação e o direito difuso à informação verdadeira**. 2º ed. Rio de Janeiro: Renovar, 2000.

CASTELLS, Manuel. **A sociedade em rede**. Trad. Roneide Vanâncio Majer. 6ª ed. – A era da informação: economia, sociedade e cultura, v. 1. São Paulo: Paz e Terra, 1992.

CELESTE, Edoardo. Digital Constitutionalism: Mapping the Constitutional Response to Digital Technology's Challenges. **HIIG Discussion Paper Series**. N. 2018-02, p. 01-24, 2018. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3219905 ou <http://dx.doi.org/10.2139/ssrn.3219905>. Acesso em: 18/02/2021.

CELESTE, Edoardo. Digital punishment: social media exclusion and the constitutionalising role of national courts, **International Review of Law, Computers & Technology**. v. 35, n. 2, p. 162-184, 2021. Disponível em: DOI: 10.1080/13600869.2021.1885106. Acesso em: 31/10/2021.

CELESTE, Edoardo. Terms of Service and Bill of Rights: New Mechanisms of Constitutionalisation in the Social Media Environment? **Internacional Review of Law, Computers & Technology**. v. 33, n. 02, p. 122-138. 2019. Disponível em: <https://doi.org/10.1080/13600869.2018.1475898>. Acesso em 26/06/2022.

CELLA, José Renato Gaziero; ROVER, Aires Jose; GOMES, Magno Federici. [Coord.]. **Direito, governança e novas tecnologias** [Recurso eletrônico on-line]. Org. CONPEDI/UFGM/FUMEC/Dom Helder Câmara. Florianópolis: CONPEDI, 2015. p. 312-327.

CHENOU, Jean-Marie; FUERTE, Juan Sebastián Rojas. The Difficult Path to the Insertion of the Global South in Internet Governance. In: OPPERMAN, Daniel. (Ed.). **Internet Governance in the Global South: History, Theory, and Contemporary Debates**. São Paulo: University of São Paulo, 2018. Disponível em: <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-65805-2>. Acesso em: 08/07/2022.

CHEQUER, Cláudio Márcio de Carvalho. **A liberdade de expressão como direito fundamental preferencial prima facie**. Rio de Janeiro: Lumen Juris, 2011.

CHEQUER, Cláudio Márcio de Carvalho. **A liberdade de expressão como direito fundamental preferencial prima facie** (análise crítica e proposta de revisão ao padrão jurisprudencial brasileiro). 2010. 340 f. Tese de Doutorado - Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2010.

CHOG, Celena. “The inventor that inspired elon musk and larry page predicted smartphones nearly 100 years ago”. Disponível em: <https://www.businessinsider.in/the-inventor-that-inspired-elon-musk-and-larry-page-predicted-smartphones-nearly-100-years-ago/articleshow/47964129.cms>. Acesso em: 25/11/2021.

COBBE, Jennifer. Algorithmic Censorship by Social Platform: Power and Resistance. **Philosophy & Techonology**. v. 34, p. 739-766, 2021. Disponível em: <https://doi.org/10.1007/s13347-020-00429-0>. Acesso em: 26/08/2021.

COMITÊ GESTOR DA INTERNET NO BRASIL. Composição do CGI.br. **CGI.br**. 2022. Disponível em: <https://www.cgi.br/membros/>. Acesso em: 01/08/2022.

COMITÊ GESTOR DA INTERNET NO BRASIL. Princípios para o Uso e a Governança da Internet no Brasil. Resolução CGI.br/RES/2009/003/P. **CGI.br**. 2009. Disponível em: www.cgi.br/resolucoes/documento/2009/003. Acesso em: 30/07/2022.

CONSELHO DE DIREITOS HUMANOS DAS NAÇÕES UNIDAS (UNHRC). (2012, 16 de julho). **A promoção, proteção e prazer dos direitos humanos na internet** (A/HRC/20/8). Disponível em: <https://digitallibrary.un.org/record/731540?ln=en>. Acesso em 30/10/2021.

CONSELHO DE DIREITOS HUMANOS DAS NAÇÕES UNIDAS (UNHRC). **Princípios Orientadores sobre Negócios e Direitos Humanos das Nações Unidas**. 2011. Disponível em: https://www.ohchr.org/sites/default/files/documents/publications/guidingprinciplesbusinesshr_en.pdf. Acesso em: 20/01/2022. p. 27-35.

CONSELHO DE DIREITOS HUMANOS DAS NAÇÕES UNIDAS (UNHRC). **Guiding Principles On Business And Human Rights At 10: Taking stock of the first decade**. A/HRC/47/39. 2021. Disponível em: <https://www.ohchr.org/sites/default/files/Documents/Issues/Business/UNGPs10/Stocktaking-reader-friendly.pdf>. Acesso em: 10/02/2022.

CONSELHO DE SUPERVISÃO. Decisão do sobre o caso 2020-004-IG-UA. **Oversightboard.com**. 2021. Disponível em: <https://www.oversightboard.com/decision/IG-7THR3SI1>. Acesso em: 10/05/2022.

COSTA NETO, João. **Liberdade de expressão: o conflito entre o legislador e o juiz constitucional**. São Paulo: Saraiva, 2017.

CRUZ, Francisco Carvalho de Brito. **Direito, democracia e cultura digital: a experiência de elaboração legislativa do Marco Civil da Internet**. 2015. Dissertação (Mestrado em Filosofia e Teoria Geral do Direito) - Faculdade de Direito, Universidade de São Paulo, São Paulo, 2015.

DAHL, Robert. **La democracia**. Uma guia para los ciudadanos. Taurus, 1999.

DAILEY, Doug. **Netezza and IBM Cloud Pak for Data: A knockout combo for tough data**. Disponível em: <https://www.ibm.com/blogs/journey-to-ai/2020/06/netezza-and-ibm-cloud-pak-a-knockout-combo-for-tough-data/>. Acesso em: 29/11/2021.

DECLARATION FOR THE FUTURE OF INTERNET. 2022. Disponível em: https://www.whitehouse.gov/wp-content/uploads/2022/04/Declaration-for-the-Future-for-the-Internet_Launch-Event-Signing-Version_FINAL.pdf. Acesso em: 30/05/2022.

DE GREGORIO, Giovanni. The rise of digital constitutionalism in the European Union, **International Journal of Constitutional Law**, v. 19, Issue 1, p. 41-70. 2021. Disponível em: <https://academic.oup.com/icon/article/19/1/41/6224442>. Acesso em 26/08/2021.

DE GREGORIO, Giovanni. Democratising online content moderation: A constitutional framework. **Computer Law & Security Review**. v. 36, p. 1-28, 2020. Disponível em: <https://doi.org/10.1016/j.clsr.2019.105374>. Acesso em: 18/07/2022.

DEMARTINI, Felipe. Youtube muda sistema de strikes e avisos por quebra de diretrizes. **Canaltech**. 2019. Disponível em: <https://canaltech.com.br/internet/youtube-muda-sistema-de-strikes-e-avisos-por-quebra-de-diretrizes-133113/>. Acesso em: 11/07/2022.

DENARDIS, Laura; MUSIANI, Francesca. Governance by Infrastructure: Introduction. In MUSIANI, Francesca; COGBURN, Derrick L.; DENARDIS, Laura; LEVINSON, Nannete S. [Orgs.] **The Turn to Infrastructure in Internet Governance. Information Technology and Global Governance**. Palgrave Macmillan, New York 2016. Disponível em: <https://ssrn.com/abstract=2730689> ou <http://dx.doi.org/10.2139/ssrn.2730689>. Acesso em: 16/06/2022.

DESNICA, Marin. SAYADCHI, Marie; KHAWJA, Umair; SZABÓ, Peter. **Google AdSense: User Modeling and Recommender Systems – Case exercise**, 2014. Disponível em: <http://htk.tlu.ee/recommendersystems/images/5/5b/Google-adsense-v1.1.pdf>. Acesso em: 18/07/2022.

DOS SANTOS, Regina Consolação et al. Fake news e a desinformação sobre COVID-19: revisão integrativa de literatura. **Research, Society and Development**, v. 11, n. 10, 2022. Disponível em: <https://rsdjournal.org/index.php/rsd/article/view/33124>. Acesso em: 25/11/2022.

DOUEK, Evelyn, COVID-19 and Social Media Content Moderation, **LAWFARE**, 2020. Disponível em: <https://www.lawfareblog.com/covid-19-and-social-media-content-moderation>. Acesso em: 20/07/2022.

DOUEK, Evelyn. Governing online speech: from “posts-as-trumps” to proportionality and probability. **Columbia Law Review**. v. 121, n. 3, p. 759-837, 2021. Disponível em: https://www.columbialawreview.org/wp-content/uploads/2021/04/Douek-Governing_Online_Speech-from_Posts_As-Trumps_To_Proportionality_And_Probability.pdf. Acesso em: 03/11/2021.

DOUEK, Evelyn. Facebook’s oversight board: move fast with stable infrastructure and humility’. **North Carolina Journal of Law & Technology**. v. 21, n. 1, p. 1-79, 2019. Disponível em: <https://scholarship.law.unc.edu/cgi/viewcontent.cgi?article=1389&context=ncjolt>. Acesso em: 03/11/2021.

DOUEK, Evelyn, The Siren Call of Content Moderation Formalism. In: BOLLIGER, Lee; GEOFFREY, Stone (Eds.). **New Technologies of Communication and the First Amendment: The Internet, Social Media and Censorship**. Oxford Press, forthcoming 2022. Disponível em: <https://ssrn.com/abstract=4005314> ou <http://dx.doi.org/10.2139/ssrn.4005314>. Acesso em: 20/07/2022.

DUCA, Danielle. **Internet in the European Union: past, present and future of digitalization**. 2021. 118 f. Dissertação (Mestrado em Estudos Europeus e Globais). Università Degli Studi di Padova. Pádua, 2021. Disponível em: https://thesis.unipd.it/bitstream/20.500.12608/32989/1/Daniele_Duca.pdf. Acesso em: 30/06/2022.

DUQUE, Marcelo Schenk. **Direito privado e constituição: drittwirkung dos direitos fundamentais**. São Paulo: Revista dos Tribunais, 2013.

DWORKIN, Ronald. **O Direito da Liberdade**: a leitura mora da constituição norte-americana. Tradução Marcelo Brandão Cipolla. 2ª ed. São Paulo: Editora WMF Martins Fontes, 2019.

DWORKIN, Ronaldo. **Uma questão de princípio**. Tradução Luís Carlos Borges. 2 ed. São Paulo: Livraria Martins Fontes Editora, 2005.

EIFERT, Martin. A Lei Alemã para a Melhoria da Aplicação da Lei nas Redes Sociais (NetzDG) e a Regulação da Plataforma. In: ABOUD, Georges; NERY JÚNIOR, Nelson; CAMPOS, Ricardo (Coord.). **Fake news e regulação**. 2 ed. São Paulo: Thomson Reuters Brasil, 2020.

EILKIN-KOREN, Niva; PEREL, Maayan. Guarding the Guardians: content moderation by online intermediaries and the Rule of Law. In: FROSIO, Giancarlo (Ed.). **The Oxford Handbook of Online Intermediary Liability**. New York. NY: Oxford University Press, 2020. Disponível em: <https://ssrn.com/abstract=3542992>. Acesso em: 28/10/2021.

EMERSON, Thomas I. Toward a general theory of the first amendment. **The Yale Law Journal**. 72:877 (1963).

ESTADOS UNIDOS DA AMÉRICA. Suprema Corte. **Abrams v. United States**, 1919.

ESTADOS UNIDOS DA AMÉRICA. **Digital Millennium Copyright Act**. Pub. L. 105-304, 112 Stat. 2860, 1998. Disponível em: <https://www.congress.gov/105/plaws/publ304/PLAW-105publ304.pdf>. Acesso em: 10/07/2022.

ESTADOS UNIDOS DA AMÉRICA. **Facebook Oversight Board**. Case Decision 2021-001-FB-FBR, OVERSIGHT BD. 9–10. May 5, 2021. Disponível em: <https://www.oversightboard.com/sr/decision/2021/001/pdf-english>. Acesso em: 01/07/2022.

FACEBOOK. **Declaração de Direitos e Responsabilidades**. 2015. Disponível em: <https://pt-br.facebook.com/legal/terms/previous>. Acesso em: 25/06/2022.

FACEBOOK. **Padrões da Comunidade do Facebook**: discurso de ódio. Disponível em: <https://transparency.fb.com/pt-br/policies/community-standards/hate-speech/>. Acesso em: 25/06/2022.

FACEBOOK. **Termos de Serviço**. 2022. Disponível em: https://www.facebook.com/legal/terms/plain_text_terms. Acesso em: 25/06/2022.

FACEBOOK. **Política de Privacidade**. 2022. Disponível em: <https://www.facebook.com/privacy/policy>. Acesso em: 20/07/2022.

FACHIN, Luiz Edson. **Direito fundamental e expressão religiosa: entre a liberdade, o preconceito e a sanção**. In: MENEZES, Joyceane Bezerra de; TEPEIDINO, Gustavo (Coord.). Autonomia privada, liberdade existencial e direitos fundamentais. Belo Horizonte: Fórum, 2019.

FARAH, André. **Liberdade de expressão e remoção de conteúdo da internet**. 2ª ed. Rio de Janeiro: Lumen Juris, 2021.

FARBER, Daniel A. **The First Amendment**. 2ª ed. New York: Foundation, 2003.

FARIAS, Edmilson. **Liberdade de expressão e comunicação: teoria e proteção constitucional**. Revista dos Tribunais, São Paulo, 2004.

FERREIRA, Felipe Grizoto. **Liberdade de expressão na era digital: desafios, perspectivas e aplicações**. Dissertação (Mestrado em Direito). Campinas: PUC-Campinas, 2021.

FISS, Owen M. **A Ironia da Liberdade de Expressão**. Rio de Janeiro: Renovar, 2005.

GILL, Lex. REDEKER, Dennis; e GASSER, Urs. Towards Digital Constitutionalism? Mapping Attempts to Craft an Internet Bill of Rights (November 9, 2015). **Berkman Center Research Publication**. N. 2015-15. Disponível em: <http://dx.doi.org/10.2139/ssrn.2687120>. Acesso em: 30/01/2021.

GILLESPIE, Tarleton. **Custodians of the internet**: Platforms, content moderation, and the hidden decisions that shape social media. Yale University Press/New Haven & London, 2018. Disponível em: https://www.researchgate.net/publication/327186182_Custodians_of_the_internet_Platforms_content_moderation_and_the_hidden_decisions_that_shape_social_media. Acesso em: 23/11/2021.

GILLESPIE, Tarleton. Content Moderation, AI, and the question of scale. **Big Data & Society**. 2020. Disponível em: <https://journals.sagepub.com/doi/full/10.1177/2053951720943234>. ou <https://doi.org/10.1177/2053951720943234>. Acesso em: 10/07/2022.

GOLDMAN, Eric. Content Moderation Remedies. **Michigan Technology Law Review**. v. 28, n. 1. p. 1-59, 2021. Disponível em: <https://repository.law.umich.edu/mltr/vol28/iss1/2>. Acesso em: 02/11/2021.

GOLDMAN, Eric; MIERS, Jess. Online Account Terminations/Content Removals and the Benefits of Internet Services Enforcing Their House Rules. **Journal of Free Speech Law**. v. 1, n. 191, Santa Clara University Legal Research Paper, 2021. Disponível em: <https://ssrn.com/abstract=3911509>. Acesso em: 03/11/2021.

GONÇALVES, Alcindo; COSTA, José A. Fontoura. **Governança Global e Regimes Internacionais**. 1 ed. São Paulo: Editora Almedina, 2011.

GONÇALVES, Géssica Brandino; RESENDE, Sarah Mota. Facebook retira rede de páginas e perfis do ar e atinge MBL. 25 de julho de 2018. **Folha de São Paulo**. Disponível em: <https://www1.folha.uol.com.br/poder/2018/07/facebook-retira-rede-de-paginas-e-perfis-do-ar-e-atinge-mbl.shtml>. Acesso em: 29/11/2021.

GORWA, Robert. What is platform governance? **Information, Communication & Society**. v. 22, n. 6, p. 854-871, 2019. Disponível em: [10.1080/1369118X.2019.1573914](https://doi.org/10.1080/1369118X.2019.1573914). Acesso em: 19/05/2022.

GORWA, Robert; BINNS, Reuben; KATZENBACH, Chistian. Algorithmic content moderation: Technical and political challenges in the automation of platform governance. **Big Data & Society**. 2020. Disponível em: <https://doi.org/10.1177/2053951719897945>. Acesso em: 20/07/2022.

GRIMMELMANN, James. The Virtues of Moderation. **The Yale Journal of Law & Technology**. v. 17, n. 1, p. 42-109, 2015. Disponível em: <https://core.ac.uk/download/pdf/72836082.pdf>. ou <https://digitalcommons.law.yale.edu/yjolt/vol17/iss1/2>. Acesso em: 10/01/2022.

HAIMAN, Franklyn S. **“Speech Acts” and the First Amendment**. Carbondale: Southern Illions University Press, 1993.

HAIMSON, Oliver L.; DELMONACO, Daniel; NIE, Peipei; WEGNER, Andrea. Disproportionate Removals and Difering Content Moderation Experiences for Conservative, Transgender, and Black Social Media Users: Marginalization and Moderation Gray Areas. **Proceeding of the ACM on Human-Computer Interacction**. v. 5, n. CSCW2, Artigo 466, p. 1-35. 2021. Disponível em: <https://doi.org/10.1145/3479610>. Acesso em: 10/07/2022.

HARFF, Graziela; DUQUE, Marcelo Schenk. Discurso de ódio nos contextos alemão e brasileiro. **A&C – Revista de Direito Administrativo & Constitucional**, Belo Horizonte, ano 21, n. 84, p. 199-225, abr./jun. 2021. Disponível em: <https://www.revistaaec.com/index.php/revistaaec/article/view/1350>. Acesso em: 19/11/2021.

HARRIS, Sukanya; JACKSON, Melody Moore; ZEAGLER, Clint. **Content Censorship on Social Media and the Impact on the Content Creator Communities**. 2022. Disponível em: https://sukanyaharris.com/wp-content/uploads/2022/05/SukanyaHarris_TermPaper.pdf. Acesso em: 15/07/2022.

HESSE, Konrad. **Elementos de Direito Constitucionais da República Federal da Alemanha**. Tradução Luis Afonso Heck. Porto Alegre: Sergio Antônio Fabris Editor, 1998.

HEY NETTO, Ricardo. **Protocolo de internet versão 6 (IPv6)**. 2018. 53 f. Monografia de Especialização em Redes de Computadores e Teleinformática, Departamento Acadêmico de Eletrônica, Universidade Tecnológica Federal do Paraná. Curitiba, 2018.

HOFMANN, Jeanette; KATZENBACH, Christian; GOLLATZ, Kirsten. Between coordination and regulation: Finding the governance in Internet governance. **New Media & Society**. v. 19, n. 9, p. 1406-1423, 2017. Disponível em: [doi:10.1177/1461444816639975](https://doi.org/10.1177/1461444816639975). Acesso em 20/05/2022.

HOFMANN, Marcia. Facebook Battle Ends in Major Victory for Users, But the War Continues. **Eletronic Frontier Foundation**. 2009. Disponível em: <https://www.eff.org/uk/deeplinks/2009/02/facebook-battle-ends-major-victory-users>. Acesso em: 20/06/2022.

HOFFMANN-RIEM, Wolfgang. **Teoria Geral do Direito Digital: transformação digital: desafios para o direito**. 2. ed. Rio de Janeiro: Forense, 2022.

HORNE, Julia Van. Shadowbanning is a Thing — and It's Hurting Trans and Disabled Advocates. **Salty**. 2020. Disponível em: <https://saltyworld.net/shadowbanning-is-a-thing-and-its-hurting-trans-and-disabled-advocates>. Acesso em: 18/07/2022.

HOWARD, Joffrey W. Free speech and hate speech. **Annual Review of Political Science**. v. 22. n. 1, p. 93-109, 2019. Disponível em: <https://doi.org/10.1146/annurev-polisci-051517-012343>. Acesso em: 20/12/2021.

HUIDOBRA, Marina Stephanie Ramos; OLIVEIRA, Natalia Rosa de. Governança global: o comitê das nações unidas para o uso pacífico do espaço (copuos) como mecanismo de governança global espacial. In: BENEVIDES, Solon Henriques de Sá; GAUDÊNCIO, Francisco de Sales (Org.). **Governança e Direitos Fundamentais: revisitando o debate entre o público e o privado**. Porto: Instituto Iberoamericano de Estudos Jurídicos, 2020.

ICANN. **Bylaws for Internet Corporation for Assigned Names and Numbers**. As amended 2 june 2022. Disponível em: <https://www.icann.org/resources/pages/governance/bylaws-en>. Acesso em 20/06/2022.

ICCPR (1966, 16 de dezembro). **Pacto internacional sobre direitos civis e políticos**. United Nations Human Rights Office of the High Commissioner. Disponível em: <https://www.ohchr.org/en/professionalinterest/pages/ccpr.aspx>. Acesso em: 30/10/2021.

ISAAC, Mike. Facebook Hampers Do-It-Yourself Mask Efforts, **New York Times**. 2020. Disponível em: <https://www.nytimes.com/2020/04/05/technology/coronavirus-facebook-masks.html>. Acesso em: 01/07/2022.

JAMART, Anne-Cleire. **Legitimacy in internet governance: Multistakeholderism and global constituent Power**. Dissertação (Mestrado em Direito). Ithaca, NY: Faculty of the Graduate School of Cornell University, 2021.

JHAVER, Shagun; APPLING, Darren Scott; FILBERT, Eric; BRUCKMAN, Amy. “Did You Suspect the Post Would be Removed?”: Understanding User Reactions to Content Removals on Reddit. **Proceedings of the ACM on Human-Computer Interaction**. Vol. 3, CSCW, Article 192. ACM, New York, NY. 2019. Disponível em: <https://doi.org/10.1145/3359294> ou <https://dl.acm.org/doi/pdf/10.1145/3359294>. Acesso em: 15/07/2022.

JHAVER, Shagun; BRUCKMAN, Amy; GILBERT, Eric. Does Transparency in Moderation Really Matter?: User Behavior After Content Removal Explanations on Reddit. **Proceedings of the ACM on Human-Computer Interaction**. Vol. 3, n. CSCW, Article 150. ACM, New York, NY. 2019. Disponível em: <https://doi.org/10.1145/3359252> ou <https://dl.acm.org/doi/pdf/10.1145/3359252>. Acesso em: 15/07/2022.

JHAVER, Shagun; GHOSHAL, Sucheta; BRUCKMAN, Amy; GILBERT, Eric. Online Harassment and Content Moderation: The Case of Blocklists. **ACM Transactions on Computer-Human Interaction**. Vol. 25, n. 2, Article 12, 2018. Disponível em: <https://shagunjhaver.com/research/articles/jhaver-2018-blocklists/jhaver-2018-blocklists.pdf>. Acesso em: 18/07/2022.

JONHSON, David R. POST, David. Law and Borders: The Rise of Law in Cyberspace. **First Monday**, v. 48 n. 1, 1996. Disponível em: <https://firstmonday.org/ojs/index.php/fm/article/view/468/824>. Acesso em: 20/10/2021.

JØRGENSEN, Rikk Frank. Human Rights and Private Actors in the Online Domain. In LAND, Molly K.; ARONSON, Jay D; (Eds.), **New Technologies for Human Rights Law and Practice**. p. 243-269. Cambridge: Cambridge University Press. 2018. Disponível em: doi:10.1017/9781316838952.011 ou https://www.cambridge.org/core/services/aop-cambridge-core/content/view/AEB97862FED32E804B078A41AFAC97E8/9781107179639c11_243-269.pdf/human-rights-and-private-actors-in-the-online-domain.pdf. Acesso em: 29/10/2021.

JØRGENSEN, Rikk Frank; PEDERSEN, Anja Møller. Online Service Providers as Human Rights Arbiters. In: TADDEO, Mariarosaia, FLORIDI, Luciano. [Eds]. The Responsibilities of Online Service Providers. **Law, Governance and Technology Series**, v. 31. p. 179-199. Springer, Cham. 2017, Disponível em: https://doi.org/10.1007/978-3-319-47852-4_10. Acesso em: 31/10/2021.

KAMARA, Seny; KNODEL, Mallory; LLANSÓ, Emma; NOJEIM, Greg; QIN, Lucy; THAKUR, Dhanaraj; VOGUS, Caitlin. Outside Looking In: Approaches to Content Moderation in End-to-End Encrypted Systems. **Center for Democracy & Technology**. 2022. Disponível em: <https://arxiv.org/ftp/arxiv/papers/2202/2202.04617.pdf> ou <https://doi.org/10.48550/arXiv.2202.04617>. Acesso em: 18/07/2022.

KANAYAMA, Rodrigo Luís. **Limites às restrições aos direitos fundamentais na atividade de investigação do Poder Legislativo**. 2007. 218 p. Dissertação (Mestrado em Direito do Estado) – Setor de Ciências Jurídicas, Universidade Federal do Paraná, Curitiba, 2007.

KANAYAMA, Ricardo Alberto. A liberdade de expressão do Marco Civil da Internet e o procedimento de notificação e retirada para as "infrações" aos direitos autorais. **Civilistica.com**. Rio de Janeiro, a. 10, n. 1, 2021. Disponível em: <http://civilistica.com/a-liberdade-de-expressao-do-marco-civil/>. Acesso em: 19/11/2021.

KELLER, Clara Iglesias. **Regulação nacional de serviços na internet: exceção, legitimidade e o papel do Estado**. Rio de Janeiro: Lumen Juris, 2019.

KELLER, Clara Iglesias. Policy by judicialisation: the institutional framework for intermediary liability in Brazil. **International Review of Law, Computers & Technology**. p. 1–19, 2020. Disponível em: doi.org/10.1080/13600869.2020.1792035. Acesso em: 01/08/2022.

KETTEMANN, Matthias C. **The Normative Order of the Internet: A Theory of Rule and Regulation Online**. Oxford: Oxford University Press. 2020. Disponível em: <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-69191-7>. Acesso em: 26/06/2021.

KETTEMANN, Matthias C.; TIEDEKE, Anna Sophia. Back up: Can users sue platforms to reinstate deleted content? **Internet Policy Review**. Alexander von Humboldt Institute for Internet and Society, Berlim, Vol. 9, n. 2, p. 1-20. 2020. Disponível em: <http://dx.doi.org/10.14763/2020.2.1484>. Acesso em: 18/07/2022.

KLONICK, Kate. The New Governors: The People, Rules and Processes Governing Online Speech. **Harvard Law Review**. v. 131. p. 1598-1670. 2018. Disponível em: https://harvardlawreview.org/wp-content/uploads/2018/04/1598-1670_Online.pdf. Acesso em: 28/05/2022.

KLONICK, Kate. The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression. **The Yale Law Journal**. v. 129, p. 2418-2499, 2020. Disponível em: https://scholarship.law.stjohns.edu/cgi/viewcontent.cgi?article=1312&context=faculty_publications. Acesso em: 20/06/2022.

KLONICK, Kate. Re-Shaming the Debate: Social Norms, Shame, and Regulation in a Internet Age. **Maryland Law Review**. v. 75. n. 4, p. 1029-1065, 2016. p. 1030. Disponível em: <https://ssrn.com/abstract=2638693> ou <http://dx.doi.org/10.2139/ssrn.2638693>. Acesso em: 20/07/2022.

KOOP, Christel; LODGE, Martin. What is regulation? An interdisciplinary concept analysis. **Regulation and Governance**. v. 11, n. 1, p. 95-108, 2017. Disponível em: http://eprints.lse.ac.uk/62135/1/Lodge_What_is_regulation_2016.pdf. ou DOI: 10.1111/reg.12094. Acesso em: 02/07/2022.

KRAUT, Robert E.; RESNICK, Paul (Ed.). **Building successful online communities: Evidence-based social design**. Cambridge, Mit Press, 2012. Disponível em: <https://www.jstor.org/stable/j.ctt5hhgvw>. Acesso em: 10/07/2022.

KUCZERAWY, Aleksandra. Responsabilidade intermediária & liberdade de expressão: Desenvolvimentos recentes na iniciativa de aviso & ação da UE. **Computer Law & Security Review: The International Journal of Technology**. v. 31, n. 1, p. 46-56, 2015. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0267364914001836>. Acesso em: 30/06/2022.

KUERBIS, Brenden; MUELLER, Milton. "The hidden standards war: economic factors affecting IPv6 deployment", **Digital Policy, Regulation and Governance**. v. 22, n. 4, p. 333-361. 2020. Disponível em: <https://doi.org/10.1108/DPRG-10-2019-0085>. Acesso em: 10/06/2022.

KURBALIJA, Jovan. **An Introduction to Internet Governance**. 7 ed. Geneva: DiploFoundation, 2016. Disponível em: https://www.diplomacy.edu/wp-content/uploads/2021/12/AnIntroductiontoIG_7th-edition.pdf. Acesso em: 20/06/2022.

LAIDLAW, Emily B. A framework for identifying Internet information gatekeepers. **International Review of Law, Computers & Technology**. v. 24, n. 3, p. 263-276, 2010. Disponível em: https://ueaeprints.uea.ac.uk/id/eprint/29193/1/A_Framework_for_Identifying_Internet_Information_Gatekeepers.pdf. Acesso em: 30/06/2022.

LAIDLAW, Emily B. Online Platform Responsibility and Human Rules. In. BELLI, Luca; ZINGALES, Nicolo. **Platform regulations: how platforms are regulated and how they regulate us**. Official outcome of the UN IGF Dynamic Coalition on Platform Responsibility. – Rio de Janeiro: Escola de Direito do Rio de Janeiro da Fundação Getúlio Vargas, 2017.

Disponível em: <https://bibliotecadigital.fgv.br/dspace/handle/10438/19402>. Acesso em: 17/05/2022.

LAIDLAW, Emily B. Online Shaming and the Right to Privacy. **Computer Science Laws**. v. 6, n. 3, p. 1-26, 2017. p. 2. Disponível em: <https://doi.org/10.3390/laws6010003> ou https://pdfs.semanticscholar.org/d08d/92023799353d3fd150896c731c0d8166ebc8.pdf?_ga=2.237340480.1368221140.1660520780-227723297.1658242439. Acesso em: 20/07/2022.

LAND, Molly K.; HAMILTON, Rebeca J. Beyond takedown: Expanding the tool kit for responding to online hate. In: DOJČINOVIĆ, Predrag (Ed.). **Propaganda and International Criminal Law: From cognition to Criminality**. 1 ed. London: Routledge, 2019. Disponível em: <https://doi.org/10.4324/978042944369>. Acesso em: 01/12/2021.

LANGVARDT, Kyle. Regulating Online Content Moderation. **The Georgetown Law Journal**. v. 106, p. 1353-1387, 2018. Disponível em: <https://www.law.georgetown.edu/georgetown-law-journal/wp-content/uploads/sites/26/2018/07/Regulating-Online-Content-Moderation.pdf>. Acesso em: 13/07/2022.

LA RUE, Frank. **Promotion and protection of the right to freedom of opinion and expression**. 2011. Disponível em: <https://www.ohchr.org/sites/default/files/Documents/Issues/Opinion/A.66.290.pdf>. Acesso em: 20/06/2022.

LAURENTIIS, Lucas Catib de; THOMAZINI, Fernanda Alonso. Liberdade de expressão: Teorias, Fundamentos e Análise de Casos. **Revista Direito e Práxis**. Rio de Janeiro, v. 11, n. 4, 2020.

LAUX, Francisco de Mesquita. Modelos de responsabilização civil por conteúdo postado na internet: Direito Comparado, Marco Civil e discussões nos tribunais superiores. **Revista de Direito Civil Contemporâneo**. v. 31. ano 9. p. 129-153. São Paulo: Ed. RT, abr./jun. 2022. p. 12. Disponível em: <http://revistadotribunais.com.br/maf/app/document?stid=st-rql&marg=DTR-2022-9860>. Acesso em: 20/07/2022.

LEE, Edward. Moderating Content Moderation: A Framework for Nonpartisanship in Online Governance. **American University Law Review**. v. 70, p. 913-1059, 2021. p. 940. Disponível em: http://aulawreview.org/au_law_review/wp-content/uploads/2021/02/Lee.to_Printer.pdf. Acesso em: 18/07/2022.

LESSIG, Lawrence. **Code**. Version 2.0. New York: Basic Books, 2006. Disponível em: https://cyber.harvard.edu/ptc2010/sites/ptc2010/images/Lessig_Code_Excerpts.pdf. Acesso em: 04/12/2021.

LESSIG, Lawrence. **Code is law: on liberty in cyberspace**. 2000 Disponível em: <https://www.havardmagazine.com//2000/01/code-is-law-html>. Acesso em: 04/12/2021.

LESSIG, Lawrence. Uma Sociedade da Informação: Livre ou Feudal. **Cúpula Mundial sobre a Informação Society (WSIS)**. 2004. Disponível em: <http://www.itu.int/wsis/docs/pc2/visionaries/lessig.pdf>. Acesso em: 04/12/2021.

LIGABO, Ambeyi; DUVE, Freimut; BERTONI, Eduardo. **International mechanisms for promoting freedom of expression**. 2011. Disponível em: www.osce.org/files/f/documents/4/0/28235.pdf. Acesso em 20/06/2022.

LEITÃO, Clara. Plataformas digitais e regulação da neutralidade da rede: como a regulação atende aos interesses de companhias com dominância de mercado. **Revista de Direito, Governança e Novas Tecnologias**. v. 7, n. 1, p. 23-39, 2021.

LEVI-FAUR, David. From “Big Government” to “Big Governance”? **Jerusalem Papers in Regulation & Governance**. n. 35, julho de 2011. Disponível em: <http://regulation.huji.ac.il/papers/jp35.pdf>. Acesso em: 10/05/2022.

LEVIN, Abigail. Pornography, Hate Speech and their challenge to Dworkin, **Public Affairs Quarterly**, v. 23, n. 4, p. 357-373, 2009.

LEVINSON, Nanette S.; MARZOUKI, Meryem. Global Internet Governance and Development: International Organizations Roles. **A 112^a Annual Meeting, American Political Science Association - APSA**. Setembro de 2017, Philadelphia (PA), United States. Disponível em: <https://hal.archives-ouvertes.fr/hal-01300442>. Acesso em: 25/10/2021.

LIGUORI FILHO, Carlos Augusto. Governança da internet no espaço regulatório global: o Idiossincrático modelo de gestão da ICANN. In: CELLA, José Renato Gaziero; ROVER, Aires Jose; GOMES, Magno Federici. [Coord.]. **Direito, governança e novas tecnologias** [Recurso eletrônico on-line]. Org. CONPEDI/UFMG/FUMEC/Dom Helder Câmara. Florianópolis: CONPEDI, p. 312-327, 2015.

LOO, Rory Van. The New Gatekeepers: Private Firms as Public Enforcers. 106 **Virginia Law Review**, 467, p. 01-49, 2020. Disponível em: https://scholarship.law.bu.edu/faculty_scholarship/800. Acesso em: 03/11/2021.

LOO, Rory Van. The Missing Regulatory State: Monitoring Business in an Age of Surveillance. **Vanderbilt Law Review**. v. 72, n. 5, p. 1563-1631, 2019. Disponível em: https://scholarship.law.bu.edu/faculty_scholarship/678. Acesso em: 03/11/2021.

LOVELUCK, Benjamin. **Redes, liberdades e controle: uma genealogia política da internet**. Petrópolis: Vozes, 2018.

MA, Renkai; KOU, Yubo. “How advertiser-friendly is my video?”: Youtuber’s Socioeconomic Interactions with Algorithmic Content Moderation. **Proceedings of the ACM on Human-Computer Interaction**, v. 5, n. CSCW2, p. 1-25, 2021. Disponível em: https://www.researchgate.net/profile/Renkai-Ma-2/publication/355078337_How_advertiser-friendly_is_my_video_Youtuber's_Socioeconomic_Interactions_with_Algorithmic_Content_Moderation/links/615c761dc04f5909fd8560f9/How-advertiser-friendly-is-my-video-Youtubers-Socioeconomic-Interactions-with-Algorithmic-Content-Moderation.pdf. ou <https://doi.org/10.1145/3479573>. Acesso em: 18/07/2022.

MACHADO, Jónatas Eduardo Mendes. **Liberdade de Expressão: dimensões constitucionais da esfera pública no sistema social**. Coimbra: Coimbra Editora, 2002.

MANTOVANI, Camila Maciel C. Alves. **Info-entretenimento na telefonia celular:** informação, mobilidade e interação social em um novo espaço de fluxos. Belo Horizonte, 2006. 152 f. Dissertação (Mestrado em Ciência da Informação) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais.

MARCUS, J. Scott; ELIXMANN, Dieter. **Regulatory Approaches to Next Generation Networks (NGNS):** An International Comparison. TPRC, 2007. Disponível em: <https://ssrn.com/abstract=2118249>. Acesso em: 05/06/2022.

MARECHAL, Natalie; RYDZAK, Jan. Meta's First Human Rights Report: The Good, the Bad, and the Missing. **Ranking Digital Rights**. 2022. Disponível em: <https://rankingdigitalrights.org/2022/07/20/metas-first-human-rights-report-the-good-the-bad-and-the-missing/>. Acesso em: 28/07/2022.

MAST, Tobias; OERMANN, Markus; SCHULZ, Wolfgang. Doing Internet Governance: Constructing Normative Structures Inside and Outside of Intermediary Organisations. **GigaNet: Global Internet Governance Academic Network**, Annual Symposium 2016. p. 3. Disponível em: <https://ssrn.com/abstract=2909367> ou <http://dx.doi.org/10.2139/ssrn.2909367>. Acesso em 01/06/2022.

MCLUHAN, Marshall. Interview with Marshall McLuhan. Entrevista concedida a L'Express. **Revista da Associação Nacional dos Programas de Pós-Graduação em Comunicação**. Brasília, v. 14, n. 3, 2011. Disponível em: <https://www.e-compos.org.br/e-compos/article/download/845/610>. Acesso em: 05/07/2022.

MEIKLEJOHN, Alexander. **Free Speech and its Relation to Self-Government**. New York: Harper & Brothers Publishers, 1948.

MENDES, Gilmar Ferreira; FERNANDES, Victor Oliveira. Constitucionalismo digital e jurisdição constitucional: uma agenda de pesquisa para o caso brasileiro. **Revista Brasileira de Direito**. vol. 16, n. 1, Passo Fundo, 2020.

MENDES, Gilmar Ferreira; FERNANDES, Victor Oliveira. Eficácia dos direitos fundamentais nas relações privadas da internet: o dilema da moderação de conteúdo em redes sociais na perspectiva comparada Brasil-Alemanha. **Revista de Direito Civil Contemporâneo**. v. 31. ano 9. p. 33-68. São Paulo: Ed. RT, abr./jun. 2022. p. 5 Disponível em: <http://revistadostribunais.com.br/maf/app/document?stid=st-rql&marg=DTR-2022-9857>. Acesso em: 05/08/2022.

MENDES, Laura Schertel. *Habeas data* e autodeterminação informativa: dois lados da mesma moeda. **Revista Direitos Fundamentais & Justiça**, Ano. 12, n. 39, p. 185-216, jul./dez. 2018.

META. **Corporate Human Rights Policy**. Disponível em: <https://about.fb.com/wp-content/uploads/2021/04/Facebooks-Corporate-Human-Rights-Policy.pdf>. Acesso em: 26/06/2022.

META. **Meta Human Rights Report: Insights and Actions 2020-2021**. Julho de 2022. p. 1-83, Disponível em: https://about.fb.com/wp-content/uploads/2022/07/Meta_Human-Rights-Report-July-2022.pdf. Acesso em: 20/07/2022.

META. **Política de Privacidade**. 2022. Disponível em: <https://www.facebook.com/privacy/policy>. Acesso em 20/06/2022.

META. Sobre a verificação de fatos no Facebook. **Central de Ajuda do Meta Business**. 2021. Disponível em: <https://www.facebook.com/business/help/2593586717571940>. Acesso em: 15/07/2022.

META. Padrões da Comunidade do Facebook. **Facebook.com**. 2022. Disponível em: <https://transparency.fb.com/pt-br/policies/community-standards/adult-nudity-sexual-activity/>. Acesso em: 15/07/2022.

MICHELMAN, Frank. Relações entre democracia e liberdade de expressão: discussão de alguns argumentos. In: SARLET, Ingo Wolfgang (org.). *Direitos fundamentais, informática e comunicação*. Porto Alegre: Livraria do Advogado, 2007.

MICKLITZ, Hans-W.; VILLANUEVA, Aurélie Anne. Responsibilites of Companies in the Algorithmic Society. In: MICKLITZ, Hans-W.; POLLICINO, Oreste, REICHMAN, Amnon, SIMONCINI, Andrea, SARTOR, Giovanni; DE GREGORIO, Gregorio. (Eds.). **Constitutional Challenges in the Algorithmic Society**. Cambridge: Cambridge University Press, 2021. Disponível em: <https://doi.org/10.1017/9781108914857>. Acesso em: 23/11/2021.

MILTON, Jonh. **Aeropagítica. Discurso pela liberdade de imprensa ao Parlamento da Inglaterra**. Rio de Janeiro: TopBooks, 1999.

MILL, John Stuart. **Sobre a liberdade**. Tradução de Ari R. Tank Brito. São Paulo: Hedras, 2010.

MIRANDA, Cássia. Facebook e Instagram excluem livre de Bolsonaro com fake news sobre aides e a vacina da covid. **O Estado de São Paulo**. Disponível em: <https://politica.estadao.com.br/noticias/geral,facebook-e-instagram-excluem-live-de-bolsonaro-com-fake-news-sobre-aids-e-a-vacina-da-covid,70003879214>. Acesso em 29/11/2021.

MONTEIRO, Artur Pericles Lima; CRUZ, Francisco Brito; SILVEIRA, Juliana Fonteles da; VALENTE, Mariana Giorgetti. **Armadilhas e caminhos na regulação da moderação de conteúdo** - Diagnósticos & Recomendações. São Paulo: InternetLab, 2021.

MORAIS, Edison A.; AMBRÓSIO, Ana Paula L. Ferramentas de busca na Internet. **Relatório Técnico: Universidade Federal de Goiás**. 2007. p. 8. Disponível em: https://ww2.inf.ufg.br/sites/default/files/uploads/relatorios-tecnicos/RT-INF_002-07.pdf. Acesso em: 15/07/2022.

MOON, Richard, **The constitutional protection of freedom of expression**. Toronto: University of Toronto Press, 2000.

MOON, Richard. The social character of freedrom of expression. **Amsterdam Law Forum**. 2:1. p. 43-48, 2009.

MOORE, Martin; TAMBINI, Damian. **Digital dominance: the power of Google, Amazon, Facebook, and Apple**. New York. NY: Oxford University Press, 2018. Disponível em: <http://14.139.206.50:8080/jspui/bitstream/1/7241/1/Digital%20Dominance.pdf>. Acesso em 27/11/2021.

MUELLER, Milton L. "A Battle for the Soul of the Internet." **Networks and States: The Global Politics of Internet Governance**, MIT Press, 2010, pp. 1-14. p. 9. Disponível em: JSTOR, www.jstor.org/stable/j.ctt5hhcz0.3. Acesso em 29/05/2022.

MUELLER, Milton. **Ruling the Root: Internet Governance and the Taming of Cyberspace**. Cambridge, MA: MIT Press, 2002.

MUELLER, Milton. The Declaration for the Future of Internet. **Blog of the Internet Governance Project (IGP)**, 29/04/2022. Disponível em: <https://www.internetgovernance.org/2022/04/29/the-declaration-for-the-future-of-the-internet/>. Acesso em 30/05/2022.

MUÑOZ, Maurício Maldonado. El principio de razoabilidad y su aplicación al estudio de validez de las normas jurídicas. **Ius Humani. Revista de Derecho**. v. 3, p. 139-174, 2012/2013. Disponível em: <https://www.iushumani.org/index.php/iushumani/article/view/40>. Acesso em 15/01/2022.

MUSIANI, Francesca. Governance algoritmica: sorveglianza, censura e diritti fondamentali. In: FOSSA, Fabio; SCHIAF-FONATI, Viola; TAMBURRINI, Guglielmo (Eds.). **Automi e persone. Introduzione all'etica dell'intelligenza artificiale e della robotica**. Carocci, p. 95-113, 2021. Disponível em: <https://halshs.archives-ouvertes.fr/halshs-03381546>. Acesso em: 23/07/2022.

MUSIANI, Francesca. La gouvernance d'Internet, brève histoire d'un espace à controverses, **Terminal**, p. 132-133, 2022, Disponível: <http://journals.openedition.org/terminal/8398>; ou DOI: <https://doi.org/10.4000/terminal.8398>. Acesso em: 25/05/2022.

MUSIANI, Francesca; SCHAFER, Valérie. Global Governance: A Short History of Debates Born With the Telegraph and Popularized by the Internet. In: BALBI, Gabriele; RIBEIRO, Nelson; SCHAFER, Valérie; SCHWARZENEGGER, Christian. [Org.]. **Digital Roots: Historicizing Media and Communication Concepts of the Digital Age**. Berlin, Boston: De Gruyter Oldenbourg, 2021. p. 117-136, Disponível em: <https://doi.org/10.1515/9783110740202-007>. Acesso em: 25/10/2021.

NAÇÕES UNIDAS. **Report of the Working Group on Internet Governance**. 2005. Disponível em: <http://www.wgig.org/docs/WGIGREPORT.pdf>. Acesso em: 22/06/2022.

NAÇÕES UNIDAS. **O Fórum de Governança da Internet**. 2015. Disponível em: <https://brasil.un.org/pt-br/71321-o-que-e-o-forum-de-governanca-da-internet-igf>. Acesso em: 25/06/2022.

NIKLAS, Jan. YouTube remove live de Bolsonaro com embaixadores. **Yahoo.com**. 2022. Disponível em: <https://esportes.yahoo.com/youtube-remove-live-bolsonaro-com-210850760.html>. Acesso em 11/08/2022.

NITRINI, Rodrigo Vidal. **Liberdade de expressão nas redes sociais: o problema da remoção de conteúdos pelas plataformas**. Belo Horizonte: Editora Dialética, 2021.

OLIVEIRA, João António França de. Rússia, **Cambridge Analytica e as Eleições Presidenciais Norte-Americanas de 2016: o ciberespaço como o mais novo domínio da conflitualidade política**. Dissertação (Mestrado). Instituto de Estudos Políticos. Universidade Católica Portuguesa. Lisboa: 2020.

OMMATI, José Emílio Medauar. **Liberdade de Expressão e o discurso do ódio na Constituição de 1988**. 5. ed. Belo Horizonte: Conhecimento Editora, 2021.

PEREL, Maayan. Digital Remedies. **Berkeley Technology Law Journal**. v. 35, n. 1, p. 1-52, 2020. Disponível em: https://lawcat.berkeley.edu/record/1174610/files/01_Perel_WEB.pdf. Acesso em: 10/07/2022.

PETERS, Anne. Compensatory Constitutionalism: The Function and Potential of Fundamental International Norms and Structures. **Leiden Journal of International Law**. v. 19, n. 3, p. 579-610, 2006. Disponível em: https://edoc.unibas.ch/5309/1/20100219153135_4b7ea0c729a55.pdf. ou doi:10.1017/S0922156506003487. Acesso em: 15/05/2022.

PETERS, Guy. Governance As Political Theory. **Oxford Handbooks On line**. Oxford University Press. 29 de março de 2012. Disponível em: <https://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199560530.001.0001/oxfordhb-9780199560530-e-2>. Acesso em: 10/05/2022.

PEREIRA, Jane Reis Gonçalves. **Interpretação Constitucional e Direitos Fundamentais**. Rio de Janeiro: Renovar, 2006.

PEREZ, Sarah. Facebook news feed changes downrank misleading health info and dangerous 'cures'. **TechCrunch**. 2019. Disponível em: <https://techcrunch.com/2019/07/02/facebook-news-feed-changes-downrank-misleading-health-info-and-dangerous-cures/>. Acesso em: 18/07/2022.

PINHEIRO, Guilherme Pereira. **Liberdade de Expressão e Neutralidade de Rede na Internet**. Rio de Janeiro: Lumen Juris, 2017.

PINHO FILHO, José Célio Belém de. **Desinformação e regulação de redes sociais digitais**. 2021. 170 f. Dissertação (Mestrado Profissional em Direito, Justiça e Desenvolvimento) Instituto Brasileiro de Ensino, Desenvolvimento e Pesquisa, Brasília, 2021.

POLETTTO, Álerton; MORAIS, Fausto Santos de. A regulação constitucional do discurso de ódio no Facebook. **Revista Direito, Inovação e Regulações**. Cascavel. Jan. 2022. v. 1, n. 1, p. 36-54, 2022.

POLLICINO, Oreste. **Digital Private Powers Exercising Public Functions: The Constitutional Paradox in the Digital Age and its Possible Solutions**. 2021. Disponível em: https://echr.coe.int/Documents/Intervention_20210415_Pollicino_Rule_of_Law_ENG.pdf. Acesso em: 27/10/2021.

POLLICINO, Oreste; DE GREGORIO, Giovanni: Shedding Light on the Darkness of Content Moderation: The First Decisions of the Facebook Oversight Board, **VerfBlog**, 2021/2/05, disponível em: <https://verfassungsblog.de/fob-constitutionalism/>, DOI: 10.17176/20210205-235959-0. Acessado em 24/11/2021.

POST, Robert. Reconciling Theory and Doctrine in First Amendment Jurisprudence. **California Law Review**. California, v. 88, p. 2353-2374, 2000.

POTKIN, Fanny. Facebook bans Myanmar military from its platforms with immediate effect. **Reuters**. 2021. Disponível em: <https://www.reuters.com/article/us-myanmar-politics-facebook-idUSKBN2AP0BK>. Acesso em: 01/07/2022.

PRINCÍPIOS DE MANILA SOBRE RESPONSABILIDADE INTERMEDIÁRIA, Versão 1.0, 2015. Disponível em: https://www.eff.org/files/2015/10/31/manila_principles_1.0.pdf. Acesso em: 12/07/2022.

PROTTI, Michel. O que você precisa saber sobre a atualização da nossa política de privacidade e em nossos termos de serviço. **Meta**. 2022. Disponível em: <https://about.fb.com/br/news/2022/05/o-que-voce-precisa-saber-sobre-a-atualizacao-da-nossa-politica-de-privacidade-e-em-nossos-termos-de-servico/>. Acesso em: 20/06/2022.

PULIDO, Carlos Bernal. **El principio de proporcionalidad y los derechos fundamentales**: el principio de proporcionalidad como critério para determinar el contenido de los derechos fundamentales vinculante para el legislador. 3ª ed. Madrid: Centro de Estudios Políticos y Constitucionales. 2007.

RANCHORDAS, Sofia, The Digitalization of Government and Digital Exclusion: Setting the Scene (29 de abril de 2020), a ser publicado em FERREIRA MENDES, G. & BLANCO DE MORAIS, C. (Org.). Direito Público e Internet: Democracia, Redes Sociais e Regulação do Ciberespaço, FGV Publicações/IDP/Univ. Lisboa, (Public Governance 4.0, 2020). University of Groningen, **Faculty of Law Research Paper**, n. 30/2020, Disponível em: <https://ssrn.com/abstract=3663051> ou <http://dx.doi.org/10.2139/ssrn.3663051>. Acesso em: 28/10/2021.

RANCHORDAS, Sofia. Digitale platformen: Globale staten naast de staat? Digital Platforms: Global States beyond the State? **University of Groningen Faculty of Law Research Paper** N. 35/2020, In Besselink, L. (red) De economische constitutie. Staatsrechtsconferentie 2019, University of Amsterdam Press, 2020. Disponível em: <https://ssrn.com/abstract=3663065>. Acesso em: 28/10/2021.

RANKING DIGITAL RIGHTS. **2015 Corporate Accountability Index**. 2015. Disponível em: <https://rankingdigitalrights.org/index2015/assets/static/download/RDR-4pager.pdf>. Acesso em: 20/06/2022

REDISH, Martin H. The Value of Free Speech. **University of Pennsylvania Law Review**, v. 130, n. 03, 1982, p. 591-645. Disponível em: <https://www.jstor.org/stable/3311836?refreqid=excelsior%3A067a6da4e7fc482a0901ef01532bd0f1>. Acesso em 13/12/2021.

RIBEIRO, Keila Cristina De Lima Alencar. **Liberdade de expressão e fake news: uma análise acerca da possibilidade de regulação das redes sociais à luz do entendimento do Supremo Tribunal Federal**. 2021. 174 f. Dissertação (Mestrado em Direito Constitucional) – Instituto Brasileiro de Ensino, Desenvolvimento e Pesquisa, Brasília, 2021.

ROBL FILHO, Ilton; SARLET, Ingo W. Estado democrático de direito e os limites da liberdade de expressão na Constituição Federal de 1988, com destaque para o problema da sua colisão com outros direitos fundamentais, em especial, com os direitos de personalidade. **Constituição, Economia e Desenvolvimento: Revista da Academia Brasileira de Direito Constitucional**. Curitiba, vol. 8, n. 14, jan.-jun. p. 112-142, 2016.

RODRIGUES JÚNIOR, Álvaro. **Liberdade de expressão e liberdade de informação: limites e formas de controle**. Curitiba: Juruá, 2009.

ROSA, Leonardo Gomes Penteado. **O liberalismo igualitário de Ronald Dworkin: o caso da liberdade de expressão**. Dissertação. 254p. Universidade de São Paulo, Faculdade de Direito, 2014.

ROSEN, Guy. An Update on Our Work to Keep People Informed and Limit Misinformation About COVID-19. **Meta**. 2020. Disponível em: <https://about.fb.com/news/2020/04/covid-19-misinfo-update/>. Acesso em: 01/07/2022.

RUSTAD, Michael L.; ONUFRIO, Maria Vittoria. Reconceptualizing consumer terms of use for a globalized knowledge economy. **University of Pennsylvania. Journal of Business Law**. v. 14, n. 4, p. 1085-1190, 2011.

SACAMAZANO, Joaquin Brage. **Los Limites a los Derechos Fundamentales**. Madrid, 2004.

SALTZ, Emily; BARARI, Soubhik; LEIBOWICZ, Claire; WARDLE, Claire. Misinformation interventions are common, divisive, and poorly understood. **Harvard Kennedy School Misinformation Review**. v. 2, n. 5, p. 1-25, 2021. p. 2. Disponível em: https://misinfoeview.hks.harvard.edu/wp-content/uploads/2021/10/saltz_misinformation_interventions_20211027.pdf. Acesso em: 18/07/2022.

SANKIEVICZ, Alexandre. **Liberdade de expressão e pluralismo: perspectivas de regulação**. São Paulo: Editora Saraiva, 2011.

SANTARÉM, Paulo Rená da Silva. **O direito achado na rede: a emergência do acesso à Internet como direito fundamental no Brasil**. 2010. 158 f., Dissertação (Mestrado em Direito, Estado e Constituição) - Universidade de Brasília, Brasília, 2010.

SANTOS, Ivo Luís Azevedo da Costa. **How does innovation influence the relationship between the socio-technical change and policy mix: acknowledging lessons from the cambridge analytica case?** Dissertação (Mestrado). Faculdade de Economia. Universidade do Porto. Porto: 2019.

SARLET, Ingo Wolfgang. Neoconstitucionalismo e influência dos direitos fundamentais no direito privado: algumas notas sobre a evolução brasileira. **civilistica.com**, v. 1, n. 1, p. 1-30,

2012. Disponível em: <https://civilistica.emnuvens.com.br/redc/article/view/29/17>. Acesso em: 15/06/2022.

SARLET, Ingo Wolfgang. **A eficácia dos direitos fundamentais**: uma teoria geral dos direitos fundamentais na perspectiva constitucional. 10 ed. 2. Tir. – Porto Alegre: Livraria do Advogado Editora, 2010.

SARLET, Ingo Wolfgang. Proteção de Dados Pessoais como Direito Fundamental na Constituição Brasileira de 1988: contributo para a construção de uma dogmática constitucionalmente adequada. **Direitos Fundamentais & Justiça**. Ano 14, n. 42, p. 179-218, jan./jun. Belo Horizonte, 2020.

SARLET, Ingo Wolfgang; HARMAN, Ivar Alberto Martins. Direitos Fundamentais e Direito Privado: a Proteção da Liberdade de Expressão nas Mídias Sociais. **Revista Direito Público**. Porto Alegre, v. 16, n. 90, p. 85-108, 2019. Disponível em: https://repositorio.pucrs.br/dspace/bitstream/10923/18863/2/Direitos_Fundamentais_e_Direito_Privado_a_Protecao_da_Liberdade_de_Expressao_nas_Mdias_Sociais.pdf. Acesso em: 24/10/2021.

SARLET, Ingo Wolfgang; MITIDIERO, Daniel; MARINONI, Luiz Guilherme. **Curso de direito constitucional**. 9 ed. São Paulo: Saraíva Educação, 2020.

SARLET, Ingo Wolfgang; SAAVEDRA, Giovani Agostini. Fundamentos Jusfilosóficos e Âmbito de Proteção de Direito Fundamental à Proteção de Dados Pessoais. **Revista de Direito Público**. Volume 17, n. 93, p. 33-57, Brasília, 2020. Disponível em: <https://www.portaldeperiodicos.idp.edu.br/direitopublico/article/view/4315>. Acesso em: 24/10/2021.

SARLET, Ingo Wolfgang; SIQUEIRA, Andressa de Bittencourt. Liberdade de expressão e seus limites numa democracia: o caso das assim chamadas “fake news” nas redes sociais em período eleitoral no Brasil. **Revista de Estudos Institucionais**. vol. 6, n. 2, p. 534-578, maio/ago, 2020.

SARLET, Ingo Wolfgang; WEINGARTNER NETO, Jayme. Liberdade de expressão: algumas ponderações em matéria penal à luz da Constituição Federal do Brasil. **Revista Espaço Jurídico**. Joaçaba, v. 18, n. 3, p. 637-660, 2017. Disponível em: <https://repositorio.pucrs.br/dspace/handle/10923/11318>. Acesso em: 8/11/2021.

SARMENTO, Daniel. **Direitos Fundamentais e Relações Privadas**. Rio de Janeiro: Lumen Juris. 2004.

SARMENTO, Daniel. A liberdade de expressão e o problema do “Hate Speech”. In SARMENTO, Daniel. **Livres e iguais**: estudos de Direito Constitucional. Rio de Janeiro: Lumen Juris, 2006.

SCHAUER, Frederick. **Free Speech**: a philosophical enquiry. Londo: Cambrigde University Press, 1984.

SHENGELIA, Nino. Constitutionalizing Role of Terms of Service of Social Media Platforms and Proposed Social Media User Bills of Rights. **International Journal of Innovative**

Technologies in Social Science. v. 7, n. 28. 2020. Disponível em: <https://rsglobal.pl/index.php/ijitss/article/view/1726/1607>. Acesso em 26/06/2021.

SHOWDEN, Edward. **Eterna Vigilância:** como montei e desvendei o maior sistema de espionagem do mundo. São Paulo: Planeta do Brasil, 2019.

SILVA, Virgílio Afonso da. **A constitucionalização do direito:** os direitos fundamentais nas relações entre particulares. São Paulo: Malheiros, 2005.

SILVESTRE, Ricardo. **Liberdade de expressão política na era digital: ameaças e soluções.** 2021. 118 p. Dissertação de Mestrado - Universidade Lusófona de Humanidades e Tecnologias - Faculdade de Ciências Sociais, Educação e Administração, Lisboa, 2021.

SIMAO, José Luiz de Almeida; RADOVALHO, Thiago. A fundamentalidade do direito à liberdade de expressão: as justificativas instrumental e constitutiva para a inclusão no catálogo dos direitos e garantias fundamentais na constituição federal de 1988. **Cadernos do Programa de Pós-Graduação em Direito PPGD/UFRGS.** Edição Digital. v. 12, n. 1, p. 203-229, 2017.

SOLUM, Lawrence B.; CHUNG, Minn. The Layers Principle: Internet Architecture and the Law. **Notre Dame Law Review.** v. 79, n. 03, p. 815-948. 2004. Disponível em: <http://scholarship.law.nd.edu/ndlr/vol79/iss3/1>. Acesso em: 22/06/2022.

SOUZA, Carlos Affonso Pereira de.; MONIZ, Pedro de Paranaguá; VIEIRA JÚNIOR, Sérgio Branco. Neutralidade da Rede, Filtragem de Conteúdo e Interesse Público: reflexões sobre o bloqueio do site Youtube no Brasil. **Revista de Direito Administrativo - RDA,** ano 16, n. 246, p. 50-78, 2007. Disponível em: <https://www.forumconhecimento.com.br/periodico/125/191/2310>. Acesso em: 08/07/2022.

SUPREMO TRIBUNAL FEDERAL. **Suspensão de Segurança 3.902/SP.** Rel. Min. Teori Zavascki. 2015.

SUSTEIN, Cass. **Democracy and the problem of free speech.** New York: The Free Press, 1993.

SUZOR, Nicolas. Digital Constitutionalism: Using the Rule to Evaluate the Legitimacy of Governance by Platforms. **Social Media + Society.** v. 4, n. 3, 2018. Disponível em: <https://journals.sagepub.com/doi/pdf/10.1177/2056305118787812>. Acesso em: 18/02/2021.

SUZOR, Nicolas. Understanding content moderation systems: new methods to understand internet governance at scale, over time, and across platforms. In: WHALEN, Ryan (Ed.) **Computational Legal Studies: The Promise and Challenge of Data-Driven Research.** Edward Elgar Publishing, p. 166-189, 2020. Disponível em: <https://eprints.qut.edu.au/129464/16/Suzor%2B2018%2BUnderstanding%2Bcontent%2Bmoderation%2Bsystems%2B-clean.pdf>. ou <https://doi.org/10.4337/9781788977456.00013>. Acesso em: 20/07/2022.

SUZOR, Nicolas; VAN GEELEN, Tess; WEST, Sarah Myers. Evaluating the Legitimacy of Platform Governance: A Review of Research and a Shared Research Agenda. **International Communication Gazette.** v. 80, n. 4, p. 385-400. 2018. Disponível em:

10.1177/1748048518757142 ou <https://osf.io/preprints/lawarxiv/jak8d/>. Acesso em: 10/06/2022.

SUZOR, Nicolas P.; WEST, Sarah Myers; QUODLING, Andrew; YORK, Jillian. What Do We Mean When We Talk About Transparency? Toward Meaningful Transparency in Commercial Content Moderation. **International Journal of Communication**. v. 13, p. 1526–1543. 2019. Disponível em: <https://ijoc.org/index.php/ijoc/article/view/9736>. Acesso em: 19/07/2022.

TEUBNER, Gunther. Horizontal Effects of Constitutional Rights in the Internet: Lergal Case on the Digital Constitution. **The Italian Law Journal**. v. 3, n. 1, p. 193-205, Julho 2017. Disponível em: <https://theitalianlawjournal.it/data/uploads/3-italj-1-2017/pdf/teubner.pdf>. Acesso em: 30/05/2022.

THE SANTA CLARA PRINCIPLES: On Transparency and Accountability in Content Moderation. Disponível em: <https://santaclaraprinciples.org/>. Acesso em: 12/07/2022.

TØNNESSON, Stein; OO. Min Zaw; AUNG, Ne Lynn. Pretending to be States: The Use of Facebook by Armed Groups in Myanmar. **Journal of Contemporary Asia**. v. 52, n. 2, p. 200-225, 2022. Disponível em: <https://www.tandfonline.com/doi/pdf/10.1080/00472336.2021.1905865>. Acesso em: 01/07/2022.

TÔRRES, Fernando Carolina. O direito fundamental à liberdade de expressão e sua extensão. **Revista de Informação Legislativa**. Ano 50, n. 200, out/dez, p. 61-80, 2013.

TRIBUNAL SUPERIOR ELEITORAL. Eleições 2022: TSE reúne-se com Google, Facebook, WhatsApp e Twitter para alinhar estratégias de combate à desinformação nas Eleições 2020. **tse.jus.br**. Brasília, 12 de nov. 2019. Disponível em: <https://www.tse.jus.br/comunicacao/noticias/2019/Novembro/tse-se-reune-com-google-facebook-whatsapp-e-twitter-para-alinhar-estrategias-de-combate-a-desinformacao-nas-eleicoes-2020>. Acesso em: 30/11/2022.

TRIBUNAL SUPERIOR ELEITORAL. Eleições 2022: TSE firma parceria com Facebook para combate à desinformação. **tse.jus.br**. Brasília, 11 de fev. de 2022. Disponível em: <https://www.tse.jus.br/comunicacao/noticias/2022/Fevereiro/eleicoes-2022-tse-firma-parceria-com-facebook-para-combate-a-desinformacao>. Acesso em: 30/11/2022.

TWITTER INC. Permanent suspension of @realDonaldTrump. **Blog.twitter**. 2021. Disponível em: https://blog.twitter.com/en_us/topics/company/2020/suspension. Acesso em: 04/11/2021.

UNITED NATIONS HUMAN RIGHTS COUNCIL. **Report 23/40 of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression**. UN. New York, 17 April 2013. A/HRC/23/40. Disponível em: <https://documents-dds-ny.un.org/doc/UNDOC/GEN/G13/133/03/PDF/G1313303.pdf?OpenElement>. Acesso em: 11/12/2021.

VALENTE, Mariana Giorgetti. A liberdade de expressão na internet: da utopia às eras das plataformas. In. FARIA, José Eduardo. [Org.] **A liberdade de expressão e as novas mídias**. 1ª ed. São Paulo: Perspectiva, 2020.

VENTURINI, Jamile; LOUZADA, Luiza; MACIEL, Marília. [et. al.]. **Termos de uso e direitos humanos**: uma análise dos contratos das plataformas online. 1 ed. Rio de Janeiro: Revan, 2019.

WALDRON, Jeremy. **The Harm in Hate Speech**, Cambridge, MA and London, England: Harvard University Press, 2012.

WAGNER, Flávio Rech; CANABARRO, Diego Rafael. A governança da Internet: definição, desafios e perspectivas. In: PIMENTA, Marcelo Soares; CANABARRO, Diego Rafael (Org.). **Governança Digital**. Porto Alegre: Editora da UFRGS/CEGOV, 2014.

WAGNER, Ben; KÜBLER, Johanne; PÍRKOVÁ, Eliška; GSENGER, Rita; FERRO, Carolina. **Reimagining Content Moderation and Safeguarding Fundamental Rights**: A Study on Community-Led Platforms. The Green/EFA in the European Parliament. 2021. Disponível em: https://www.greens-efa.eu/files/assets/docs/alternative_content_web.pdf. Acesso em: 15/07/2022.

WEST, Sarah Myers. Censored, suspended, shadowbanned: User interpretations of content moderation on social media platforms. **New Media & Society**. v. 20, n. 11, p. 4366-4383, 2018. Disponível em: <https://doi.org/10.1177/1461444818773059>. Acesso em: 19/07/2022.

WORLD SUMMIT ON THE INFORMATION SOCIETY. **Tunis agenda for the information Society**. 2005. Disponível em: <https://www.itu.int/net/wsis/docs2/tunis/off/6rev1.html>. Acesso em: 15/05/2022.

WYLIE, Christopher, *apud* GUIMÓN, Pablo. **O ‘Brexit’ não teria acontecido sem a Cambridge Analytica**. Disponível em: https://brasil.elpais.com/brasil/2018/03/26/internacional/1522058765_703094.html. Acesso em: 30/11/2021.

YANG, Kai-Cheng; PIERRI, Francesco; HUI, Pik-Mai [et al.]. The COVID-19 Infodemic: Twitter versus Facebook. **Big Data & Society**, 2021. Disponível em: <https://doi.org/10.1177/20539517211013861>. Acesso em: 21/07/2022.

ZITTRAIN, Jonathan. **Three Eras of Digital Governance**. Disponível em: <https://ssrn.com/abstract=3458435> ou <http://dx.doi.org/10.2139/ssrn.3458435>. Acesso em: 08/07/2022.

ZUCKERBERG, Mark. Post Announcing Indefinite Suspension of Trump’s Accounts. **Facebook**. Disponível em: <https://www.facebook.com/zuck/posts/10112681480907401>. Acesso em: 06/11/2021.

ANEXO I**PRIMEIRO RELATÓRIO SOBRE DIREITOS HUMANOS DA EMPRESA META - 2022**

JULY 2022



Meta Human Rights Report

Insights and Actions 2020 - 2021

About this Report

In March of 2021, Meta adopted its [Human Rights Policy](#) in which we commit to reporting annually on how we are addressing our human rights impacts, including relevant insights arising from human rights due diligence, and the actions we are taking in response. This is our first annual report, covering our learnings and progress from January 1, 2020 through December 31, 2021.

The scope of this report is Meta Platforms, Inc. (formerly known as Facebook, Inc.) and our assessment of what we consider to be

the company's salient human rights risks, defined by the scale, scope, irremediable character and likelihood of impact. Our salience assessments are complemented by an additional materiality assessment included in [Meta's 2021 Sustainability Report](#).

This report is inspired by Principle 15 of the UN Guiding Principles on Business and Human Rights which makes it clear that companies must “know and show” that they respect human rights.

Note on Terminology

On October 28, 2021, Facebook, Inc. changed its name to Meta Platforms, Inc. For consistency, this report uses “Meta” to refer to the company both before and after October 28, 2021. References to “Facebook” apply only to the social media platform, not the company as a whole.

Meta's Corporate Human Rights Policy applies enterprise-wide. Each Meta product and entity has its own policies and procedures with, at times, varying human rights impacts. This report references a particular policy or procedure that may not apply enterprise-wide (e.g., the Community Standards, which apply only to Facebook). Further, this report references actions taken by Meta as a company regarding a specific entity. Such a statement is not intended to imply that Meta took that action regarding all entities. No statement in this report is intended to create — or should be construed as creating — new obligations (legal or otherwise) regarding

the application of a policy or procedure to other products or entities. If a policy is labeled a “Facebook” policy, it does not apply to other entities, e.g., Instagram or WhatsApp.

For example (and in contrast to other Meta technologies), WhatsApp is an end-to-end encrypted messaging and calling application with unique human rights touchpoints. The report considers the human rights impacts of end-to-end encryption. Likewise, this report's discussion of content moderation and related issues on Facebook and Instagram do not apply to WhatsApp. Unless a policy is specified as applying to WhatsApp, it does not apply to WhatsApp. Further, while many content moderation decisions described in this report apply to Instagram and Facebook, there are intentional distinctions in policies and procedures between the platforms. Unless a policy is specified as applying to Instagram, it does not apply to Instagram.

Executive Summary

Background

Born out of the despair of war, international human rights laws and treaties represent humanity's best hopes for itself.

Meta's mission is to give people the power to build community and bring the world closer together. We seek to advance our mission in a manner consistent with the promotion of humanity's fundamental rights: freedom of expression, privacy, non-discrimination, and more.

Meta recognizes its potential to impact human rights and has taken steps to demonstrate our commitment to human rights principles. We joined the Global Network Initiative, and made commitments in our Corporate Human Rights Policy to respect human rights as set out in the United Nations Guiding Principles on Business and Human Rights (UNGP) and other international human rights standards. This report provides an update on our progress.



Theory of Change

Our human rights policy helps ensure that Meta incorporates consideration of human rights principles into our decisions and actions and supports the evolution of the business and

human rights agenda. Our strategy leverages our mission and values. Simply put, we seek to translate human rights guidance into meaningful action, every day.



Salient Risks

Under the [UNGP](#) framework, we conducted due diligence to identify the most salient risks to human rights that could be associated with the

company's products. In this report, we identify those risks and share the steps we have taken to help manage them.



Right to Freedom of Opinion and Expression

Meta's platforms advance freedom of expression — which is interdependent with all fundamental human rights. For example, freedom of expression requires security of the body, protection from non-discrimination, and privacy. These rights are not in tension; one secures the other. We manage risks related to balancing freedom of expression and other fundamental rights on our social media platforms through our Facebook [Community Standards](#) and Instagram [Community Guidelines](#) (see, the [Community Standards Enforcement Reports](#)).

During the reporting period, we updated our policies as part of our continuing effort to balance these rights in an increasingly dynamic world: including, [specifically referencing human rights principles](#), [clarifying our health misinformation policies](#), [enhancing our bullying and harassment policy to include stronger protections against gender-based harassment](#) for everyone including public figures, enacting prohibitions against certain [mass harassment or brigading](#), and [expanding our policies that prohibit veiled and implicit threats](#), among others. Meta further provides users with a path to disagree with content policy decisions, as described in greater detail in this report under [Providing Access to Remedy](#).



Right to Privacy

The right to privacy is essential to realizing the right to freedom of expression. On [WhatsApp](#), and [expanding to our other messaging apps](#), we offer [end-to-end encryption](#), an [important privacy and security function](#) to protect people's private messages and calls, including journalists and human rights defenders.

We protect users from unlawful or overbroad government data requests via our Meta [Data Policy](#), our [WhatsApp Privacy Policy](#), our dedicated Law Enforcement Response Team, and a UNGP-informed tool assessing our responsiveness to data requests from law enforcement authorities (see, [Government Requests for User Data Transparency Reports](#)). There are many legitimate and lawful reasons why governments request

information from social media companies: international criminal investigations, anti-terrorism work, prevention of child sexual abuse and other efforts intended to protect the lives, security and rights of people in their countries. We respond to such requests for information that are consistent with internationally recognized standards on human rights, including due process, and the rule of law. If we determine that a government request is not consistent with applicable law or our policies, or unlawful (for example, overly broad, or legally deficient in any way), we do not provide governments with direct access or “back doors” to people’s information and we would challenge any order that sought to have us redesign our systems to undermine the encryption we provide to protect people’s data.



Right to Life, Liberty and Security of Person

Our Facebook [Community Standards](#) and Instagram [Community Guidelines](#) seek to protect freedom of expression and the salient risks to freedom of life, liberty, and security of the person, and include rules against [Violence and Incitement](#), [Dangerous Individuals and Organizations](#), [Coordinating Harm and Promoting Crime](#), [Bullying and Harassment](#), [Human Exploitation](#), [Hate Speech](#), [Violent and Graphic Content](#), and [Misinformation](#). We manage risks of abuse of our products for human trafficking and exploitation via in-product features to raise awareness of trafficking, deter violating behavior, and offer support to victims. We are working to increase our ability to identify the illicit actors, networks, organizations, and businesses that perpetrate these activities and disrupt them accordingly.

We manage human trafficking and exploitation risks in our supply chain through our [Code of Conduct](#), our [Responsible Supply Chain](#) program, and our membership in the [Responsible Business Alliance](#) and the [Responsible Labor Initiative](#) (see, [Anti-Slavery and Human Trafficking Statement](#)).

Moreover, as we detail in this report, [protecting Human Rights Defenders is a core pillar of our Human Rights Policy](#).



Rights to Equality and Non-Discrimination

Our Community Standards and Community Guidelines address [hate speech](#), and we have [advertising policies on non-discrimination](#). We have dedicated Civil Rights and Human Rights teams; and we have cross-company initiatives for responsible innovation and accessibility.



Best Interests of the Child

In developing youth-focused features, we are listening to [leading experts on child safety](#) (including the [National Center for Missing and Exploited Children](#)). We published a [Parents Guide](#) and other features to [increase child and teen safety](#) on Instagram, and continued our work to fight child exploitation on [WhatsApp](#) and [Facebook and Instagram](#).



Right to Public Participation, to Vote, and to be Elected

Meta is committed to protecting the right to public participation, giving everyone a voice through our products, and empowering people to vote. Over the last few years, Meta [has made significant investments](#) in teams and technologies to better protect free and fair elections, including dedicated teams focused on election integrity and products that bring relevant and reliable [voting information](#) to people. We [continue to learn](#) how to treat different types of speech that have the most meaningful impact on public opinion and how people vote at the polls, such as ads that discuss, debate, or advocate for or against important topics under our [Ads about Social Issues, Elections or Politics Policy](#). We fundamentally believe in freedom of expression, while recognizing the need to set guidelines for ads about elections and to adopt measures to reduce the spread of harmful misinformation.

Meta's Human Rights Work In Practice

I. Applying Relevant Policies;

Meta's respect for human rights — and their underlying principles of equality, safety, dignity, privacy and voice — is applied through existing policies across our company, including the Facebook [Community Standards](#), Instagram [Community Guidelines](#), WhatsApp [Terms of Service](#), the [Responsible Supply Chain](#) program, Facebook's [Privacy Principles](#), our [Law Enforcement Guidelines](#), our [Data Policy](#), [Responsible AI pillars](#), and our [Diversity and Inclusion](#) practices, among other policies.



II. Conducting Human Rights Due Diligence and Disclosure;

We endeavor to assess human rights impacts that may be linked directly to Meta, take appropriate action and implement findings, monitor implementation, and report annually on “relevant insights” and “actions” under UNGP 21 as defined in our [Corporate Human Rights Policy](#). From 2020 — 2021, our UNGP-informed diligence work included:

- Detailed human rights impact assessments, e.g., [the Philippines](#) and [End-to-End Encryption](#);
- Human rights due diligence on rapidly evolving situations, e.g., Israel-Palestine (to be published in Q3 2022 in response to the Oversight Board), and product initiatives, e.g., [Ray-Ban Stories](#);
- Integrated approaches, embedding UNGP prioritization criteria, and well-defined sets of human rights standards that allow us to assess and to identify human rights concerns for countries at risk and for particularly vulnerable populations.



III. Providing Remedies for Human Rights Impacts

Users can challenge some decisions regarding removal of their account or content via [the help center](#), and appeal up to the [Oversight Board](#). A [detailed human rights report](#) guided the creation of the Oversight Board. The Oversight Board [consists](#) of rights experts from a variety of cultural and professional backgrounds. The Board's purpose is to: "[p]rotect free expression by making principled, independent decisions about important pieces of content and by issuing policy advisory opinions on [Meta's] content policies" (See, [Oversight Board Transparency Reports](#)).



IV. Protecting Human Rights Defenders

We are building on our existing work to protect human rights defenders and their accounts. Efforts include: [combatting advanced threat actors](#) targeting them, protecting them from incorrect content removals using [Cross Check](#), offering advanced [security options](#), taking steps to thwart unauthorized access to the accounts of defenders who are arrested or detained, and partnering with human rights organizations on outreach and training. In November 2021, we launched the [Human Rights Defender Fund and Journalist Safety Initiative](#), in partnership with [Civil Rights Defenders](#) and [International Center for Journalists](#), which gives offline assistance to human rights defenders facing critical threats and supports new digital security efforts.



V. Governance, Oversight and Accountability

Meta created a global Human Rights Policy Team and tasked it with guiding the implementation of the Human Rights Policy, training stakeholders on the Policy, and applying the Policy throughout the company by integrating into existing and developing policies, programs, and products. The Team is accountable to the [Audit Risk and Oversight Committee](#) of the Board of Directors. Meta's [Human Rights Policy](#) and its commitments to protect human rights are enterprise-wide.



The Road Ahead

Our Human Rights Policy laid a strong basis for action. In the coming years, we intend to build the maturity of our program, function, and reporting.

We continue to strive to better protect rights in countries in conflict and crisis, and for particularly disadvantaged groups. We understand that our work must address today's technologies and look towards the future.

Human and civil rights must be at the center of our development of the [metaverse](#).

Protecting rights in the digital space is one of the existential challenges for the human rights movement and for all who care about human rights. In the face of this, we approach our work with humility and determination.



Table of Contents

Introduction	13
Human Rights and Policy Timeline	14
Part 1: Meta's Human Rights Commitments	16
Background	17
The Global Network Initiative	18
United Nations Guiding Principles on Business and Human Rights	18
Our Human Rights Vision & Strategy: The Meta Human Rights Policy and Our Theory of Change	19
From "Know and Show" to "Show, Not Tell"	20
Our Approach to Due Diligence	24
Meta's Salient Risks	25
How Meta identifies its Salient Risks	26
Salient Risks	27
Part 2: Meta's Human Rights Policy in Practice	28
Applying Relevant Policies	29
Content Policies	30
Actions with Regard to Salient Risks	31
Community Standards in Focus: Dangerous Organizations	34
Privacy	35
Privacy Program Overview	35
Salient Risks in Privacy	38
Protecting Users from Unlawful or Overbroad Government Data Requests	39
Properly Collecting, Using, Storing, and Deleting User Data	41
Civil Rights	42
Responsible Innovation and Responsible AI	44
Accessibility	46
Anti-Slavery and Human Trafficking	47
Election Integrity	48
United States Presidential Election 2020	48
Myanmar Election 2020	49
Preparation for Philippines Presidential Election 2022	50
COVID-19	51
Emergency Due Diligence	52
Content Policy Due Diligence	52
Privacy Due Diligence and Data for Good	53

Conducting Human Rights Due Diligence and Disclosure	54
Human Rights Impact Assessments	55
Philippines HRIA	56
Summary Disclosure - India Human Rights Impact Assessment	57
Product-Focused Human Rights Impact Assessments	61
End-to-End Encryption (“E2EE”)	61
Due Diligence Exercises	63
Ray-Ban Stories	64
Integrated Analytical Approaches	65
Countries At Risk	65
Our Program in Action: Case Studies	67
Myanmar	67
Ethiopia	68
Security Policies	69
Combating Emerging Threat Actors	70
Protecting Users Against Espionage	71
Providing Access to Remedy	72
The Oversight Board	73
Code of Conduct	74
Protecting Human Rights Defenders	75
Governance, Oversight, Accountability	77
Governance Model	77
Stakeholder Engagement	79
Program Spotlight: Trusted Partners	81
Engagement with International Organizations	82
A Final Note	83

Introduction

While technology has changed much about our world, the principles of human rights remain timeless. As the [Universal Declaration of Human Rights](#) makes clear: “All human beings are born free and equal in dignity and rights” and “everyone has the right to freedom of opinion and expression.”

When first drafted, these principles were meant to guide the behavior of governments. Today, they reach far beyond, influencing the policies and behaviors of global companies like Meta, whose decisions can make a real difference in how people’s rights are exercised, protected and respected.

Indeed, the digital sphere represents a unique challenge and opportunity for the human rights movement. Companies, the United Nations, governments, and civil society must work together to protect basic human rights online, so that as we continue to advance and innovate, we ensure human rights are respected at the frontier of new technologies.

To protect human rights online, there are rarely simple answers, only a careful balance to be struck between competing values. With billions of people using Meta’s apps and services across the world every day, it is incumbent on us to have the right policies in place, and the right processes for acting on them.

Last year, we launched our first [Corporate Human Rights Policy](#), which formalizes the company’s commitment to human rights and explains how we apply its principles to our platforms, products, policies, transparency, and programming efforts.

Under this policy, we committed to publishing an annual human rights report that would keep the public informed of our progress.

This report is inspired by Principle 15 of the UN Guiding Principles on Business and Human Rights which makes it clear that companies must “know and show” that they respect human rights.

It sums up our progress over the past two years, and explains how we think about our human rights responsibilities and what we do to uphold them. We seek to embed our commitments in a governance model which supports integration of our human rights work with ongoing activities and policies on civil rights and Environmental, Social and Corporate Governance (ESG) efforts, as part of the company’s culture, governance, decision-making processes and communication strategies.

As a next step, we’ll build on preliminary work to carry out a comprehensive human rights risk (also known as a salient risk) assessment to help prioritize our approaches and align our resources and commitments.

Most of all, we strive for our efforts to make a difference, and for our human rights commitments to be reflected in today’s technologies and in the metaverse to come.

Human Rights and Policy Timeline

2011

Adoption of the [UN Guiding Principles on Business and Human Rights](#) by the UN Human Rights Council



2016

WhatsApp completes its [rollout of end-to-end encryption](#)

2019

Creation of Meta Human Rights Policy Team

2020 May.

Announcement of [first 20 members of the Oversight Board](#)



2020 July.

Release of [Civil Rights Audit Report](#)

2020 Oct.

Oversight Board [begins operations](#)

2013

Meta [joins the Global Network Initiative](#)



2018

Public release of [Myanmar independent human rights impact assessment](#)

2020

Voluntary disclosure of data to the [UN mechanism](#) investigating events in Myanmar begins

2020 May.

Public release of executive summaries and Meta responses to human rights impact assessments for [Cambodia](#), [Indonesia](#), and [Sri Lanka](#)



2021 March.

Launch of the Meta
Human Rights Policy



2021 Sep.

Meta joins the
UN Global Compact



2021 Nov.

Publication of the
Civil Rights Audit
progress report



2022 April.

Publication of End-to-End
Encryption HRIA

2021 Jan.

Meta appoints first VP
of Civil Rights



2021 June.

Launch of revised
Meta Code of Conduct

2021 Oct.

Launch of revised content
policies for public figures and
human rights defenders



2021 Nov.

Launch of pilot human rights
defender fund and journalist
safety initiative in APAC

2021 Dec.

Launch of Philippines
Human Rights Impact
Assessment



PART 1:

Meta's Human Rights Commitments

Background

In April of 2004 — two months after Meta’s initial founding — Shi Tao, a Chinese journalist, logged into his Yahoo! account and emailed a document about the 15th anniversary of the 1989 Tiananmen Square protests to Democracy Forum, a US-hosted website. Intense [controversy](#) erupted when it was revealed that Yahoo!, a US-based internet company, acceded to requests by the Chinese government for Shi Tao’s personal data. Tao was convicted of sharing state secrets, and sentenced to 10 years in prison.

Following this incident, [information and communications technology companies and their human rights stakeholders aligned on a framework](#) to protect people’s rights to freedom of expression and privacy against overbroad or unlawful government requests; essentially, building

a way to encourage rights-respecting practices by telecommunication and technology companies.

In 2008, this effort became the [Global Network Initiative](#) (“GNI”). Directly interpreting the relevant provisions of the [International Covenant on Civil and Political Rights](#), its members committed to implementing certain human rights protections in their decision making and processes. It was — and is — an important innovation in the field of business and human rights. Meta became a GNI member in 2013.

Building on GNI frameworks, Meta [adopted](#) its [Corporate Human Rights Policy](#), and [created a policy team](#) that works to integrate human rights principles across the company.



A. The Global Network Initiative

Global Network Initiative members must commit to its Principles on Freedom of Expression and Privacy (“GNI Principles”). These Principles are based on internationally recognized laws and standards and are intended to “provide direction and guidance to the ICT [Information and Communications Technology] industry and its stakeholders in protecting and advancing the enjoyment of these human rights globally.”

Meta joined the GNI in 2013, recognizing how “advancing human rights, including freedom of expression and the right to communicate freely, is core to our mission” and that by joining, we hoped to “shed a spotlight on government practices that threaten the economic, social and political benefits the internet provides.”

To assess whether GNI member companies are “making good-faith efforts to implement the GNI Principles with improvement over time,”

member companies are required to undergo periodic independent assessments conducted by accredited third-party assessors. GNI members are accountable for identifying, preventing, and mitigating risks of overbroad or legally deficient government requests to restrict freedom of expression or violate users’ right to privacy.

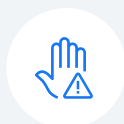
Since joining the GNI, Meta has cooperated fully with two assessments, an assessment in 2015—2016, and an assessment in 2018—2019, the latter of which found Meta is making “good-faith efforts to implement the GNI Principles with improvement over time” and commended Meta for “strengthen[ing] its systematic review of both privacy and freedom of expression.” Meta’s next GNI assessment is ongoing as of this report’s publication.

Further information regarding the GNI’s mission and principles can be found on the GNI website.

United Nations Guiding Principles on Business and Human Rights

In 2011, the United Nations Human Rights Council unanimously adopted the United Nations Guiding Principles on Business and Human Rights (“UNGPs”).

Historically, governments were obliged to protect human rights. The UNGPs: changed this by articulating a role for businesses in the rights framework. For example, under UNGP II(A)(13), businesses have a responsibility to:



- a. Avoid causing or contributing to adverse human rights impacts through their own activities, and address such impacts when they occur; and



- b. Seek to prevent or mitigate adverse human rights impacts that are directly linked to their operations, products or services by their business relationships, even if they have not contributed to those impacts.

Unlike treaty law, the UNGPs are voluntary and non-binding. Companies seeking to follow the UNGPs signal that commitment in a clear and public policy statement.

B. Our Human Rights Vision & Strategy: The Meta Human Rights Policy and Our Theory of Change

Meta’s mission — to give people the power to build community and bring the world closer together — is strongly aligned with human rights principles. Our mission guides how we work, what we prioritize, and which principles we use to make decisions. In this sense, our mission is inherently, closely associated with the values and rights of the Universal

Declaration of Human Rights, as well as the many treaties, practices, and norms that succeeded it. We strive for a world where businesses respect rights, governments protect rights, and people prosper.

Simply put — we seek to translate human rights guidance into meaningful action, every day.



From “Know and Show” to “Show, Not Tell”

[UNGP 21](#) states that companies should have in place policies and processes by which they can “know and show” their human rights responsibilities, including transparent and accountable communication to people or groups whose rights may be impacted by a company’s operations.

With this in mind, and in pursuit of meaningful transparency, we worked to build an ethos of human rights action that is grounded in “show, not tell.” e.g.:

- Newsroom Posts from the Global Human Rights Policy Team on human rights issues, including [Our Commitment to Human Rights](#), [Our Approach to Countries at Risk](#).
- Our [Transparency Center](#) provides a hub for Facebook’s and Instagram’s integrity and transparency work, acting as a central destination for all updates on how we [enforce Facebook’s Community Standards](#) and how we respond to decisions, recommendations, and case updates from the Oversight Board.
 - We published [an end-of-year threat report](#) on six adversarial networks we found and removed for Coordinated Inauthentic Behavior, Brigading and Mass Reporting.
 - We publish a [quarterly Community Standards Enforcement Report](#) covering Facebook and Instagram.
 - We semi-annually publish data on how we enforce our [Intellectual Property](#) policies.
 - We regularly publish data on [Government Requests for User Data](#).
 - We semi-annually publish data on [Content Restrictions Based on Local Law](#).
 - We semi-annually publish data on [Internet Disruptions](#).
 - We report quarterly on [Widely Viewed Content Report: What People See on Facebook](#).
- We publish reports that apply human rights principles:
 - [Conflict Minerals Policy and Annual Report](#)
 - [Annual Diversity Report](#)
 - [Annual Sustainability Report](#)
 - [Privacy Progress Update](#)
 - [Progress on Civil Rights Audit Commitments](#)
 - [Anti-Slavery and Human Trafficking Statement](#)



Defining What Matters: Our Human Rights Policy

Building on our mission, we adopted our [Corporate Human Rights Policy](#) on March 16, 2021, committing ourselves to respecting human rights as set out in the UNGPs and other human rights

standards.¹ The Policy guides teams to build rights-respecting products, respond to emerging crises, and work with speed and agility to embed human rights at scale. Also noteworthy:

We call the policy a *policy* — not a “statement.” It is an enterprise-wide commitment; part of our governance, not just a reflection of a position.



- Our [Code of Conduct](#) requires personnel to respect the Policy.
- Meta recognizes an expansive set of relevant human rights standards and instruments — not just freedom of expression — and other rights contained in the International Covenant on Civil and Political Rights (ICCPR).² In particular, the policy is framed to recognize the importance of non-discrimination, equity, and equality — and links to the work of the Meta Civil Rights team.

Section Two (“How We Implement”) describes the Policy’s five main pillars and includes a commitment to regularly reviewing and updating the Policy.

01. Meta joins other technology companies in making human rights responsibilities more explicit. For example, Apple published its [human rights commitment](#) in August 2020, the same year that Google amended the charter of their [Audit and Compliance Committee](#) to include oversight of civil and human rights issues.
02. E.g. the [United Nations Guiding Principles on Business and Human Rights](#), which encompasses internationally recognized human rights as defined by the International Bill of Human Rights, including of the Universal Declaration of Human Rights; the International Covenant on Civil and Political Rights; and the International Covenant on Economic, Social and Cultural Rights — as well as the [International Labour Organization Declaration on Fundamental Principles and Rights at Work](#).



Our commitments: In the Human Rights Policy, we recognize:

- [The United Nations Guiding Principles on Business and Human Rights \(UNGPs\)](#)
- [The International Bill of Human Rights \(the Universal Declaration of Human Rights; the International Covenant on Civil and Political Rights; the International Covenant on Economic, Social and Cultural Rights\)](#);
- [The International Labour Organization Declaration on Fundamental Principles and Rights at Work](#);
- [The International Convention on the Elimination of All Forms of Racial Discrimination](#);
- [The Convention on the Elimination of All Forms of Discrimination Against Women](#);
- [The Convention on the Rights of the Child](#);
- [The Convention on the Rights of Persons with Disabilities](#);
- [The Charter of Fundamental Rights of the European Union](#);
- [The American Convention on Human Rights](#);
- [The Global Network Initiative \(GNI\) Principles on Freedom of Expression and Privacy, and their associated Implementation Guidelines](#);
- [The OECD Principles on Artificial Intelligence](#); and
- [The UN Declaration on Human Rights Defenders](#).

We recognize that the universal obligation of non-discrimination is a necessary — but not sufficient — condition for real, lived, equality.

We strive to respect domestic laws. When faced with conflicts between such laws and our human rights commitments, we seek to honor the principles of internationally recognized human rights to the greatest extent possible. In these circumstances we seek to promote international human rights standards by engaging with governments, and by collaborating with other stakeholders and companies.

Meta is anchored in the value “Live in the Future”

Bringing the Policy to Life: A Story of Culture and Progress

Meta is anchored in the value “[Live in the Future](#),” where we build the future of distributed work and where opportunity isn’t limited by geography; to do this, we develop products using agile software development practices.³

Our human rights strategy is built from a well-known theory of change, the Knowledge, Attitudes, and Practices Model (KAP). Researchers assume that knowledge, attitudes, and practice are related, and that knowledge and attitudes directly influence practice.⁴

We are a mission-driven company where employees are typically aligned with human rights norms. In turn, this consensus leads to a company-wide community that wants to protect and advance human rights. Our Policy and integrated analytical approaches provide guidance, expectations and guardrails for employees to apply in their work.

Human rights defenders, subject matter experts, external stakeholders, and our Human Rights Policy Team contribute to our **knowledge**. The Human Rights Policy Team shares that knowledge by advising teams and decision-makers; participating in external fora; developing role-specific and enterprise-wide training; and by performing useful and durable human rights due diligence.

Together, we create rights-respecting **practices**, decisions, approaches, and products.

To work at scale, we are guided by UNGP prioritization criteria, and look for opportunities to affect a broader system that touches multiple products and countries. By way of example, the Meta-created Oversight Board’s role in content policy development and feedback advances human rights as well as other content moderation imperatives. Likewise, so does our workstream focused on mitigating risks of harm in countries experiencing or at risk of violence.

Using this theory of change, we are building an impactful and durable human rights risk management system; one that has greater strength because it is closely aligned with our mission.

03. According to the [Agile Alliance](#), agile software development is “an umbrella term for a set of frameworks and practices based on the values and principles expressed in the [Manifesto for Agile Software Development](#) and the [12 Principles](#) behind it.” It refers to an environment where “Solutions evolve through collaboration between self-organizing cross-functional teams utilizing the appropriate practices for their context.” For an overview of agile and relevance across a broad range of industries, see Darrell Rigby, Jeff Sutherland, and Hirotaka Takeuchi “Embracing Agile: How to Master the Process that’s Transforming Management” [Harvard Business Review](#), May 2016. See also, [Nikola Jurisic](#), [Michael Lurie](#), [Philippine Risch](#), and [Olli Salo](#): [Doing vs Being: Practical Lessons on Building an Agile Culture](#). McKinsey and Company, August 4 2020

04. Originally pioneered in public health research, “K” stands for knowledge of a problem or disease, “A” for attitude towards a problem or disease, and “P” for practice or preventive behavior to protect against the problem or disease. A relevant bibliography is cited in T.H. Wan et al., [A Systematic Review of KAP-O Framework for Diabetes](#), [ESMED Medical Research Archives Vol 3 No 9 \(2016\): April, Vol.3 Issue 9](#)

C. Our Approach to Due Diligence

[UNGP 17](#) states that companies should carry out human rights due diligence in order to “identify, prevent, mitigate and account for how they address their human rights impacts.” [UNGP 24](#) guides companies to prioritize their actions to address actual and potential human rights impacts according to severity.

We based our early approaches to human rights due diligence on our obligations as members of the GNI; specifically, we put into place strong operational, policy, and legal measures designed to scrutinize governmental takedown requests and requests for personal data of users.

We expanded our approach with the November 2018 release of an independent human rights impact assessment (“HRIA”) of Facebook’s impact in [Myanmar](#).⁵

Methodologies for human rights due diligence are evolving. HRIAs are only one form of diverse human rights due diligence offerings. Historically, many HRIAs stemmed from the extractive or natural resource industries and focused on a specific defined physical location, or on supply chain issues. Today, initiatives such as the [UN B-Tech](#) project seek to build on these methodologies to better fit risks faced by technology companies.

Global frameworks agree that the defining element of a good faith HRIA is that it involves direct consultation with affected rights-holders, as made clear in [UNGP 21](#) which requires companies to have in place policies and processes by which they can “know and show” their human rights responsibilities, including transparent and accountable communication to people or groups whose rights may be impacted by a company’s operations.⁶ Since 2018, Meta has rapidly experimented with ways to “show, not tell” our commitment to human rights risk management through our due diligence program. Building on country-based HRIAs from 2018 and 2019, Meta tested a variety of due diligence approaches

during this reporting period. Our goal has been to turn insights into actionable approaches that work for a company of Meta’s size, scale, and the culture of agile software development.

In addition, we engaged in crisis support; continuous due diligence processes, where we focus on a key challenge for an ongoing period; as well as ongoing content policy feedback, and real time decision-making support.

We also sought to be an industry leader in assessment and transparency. On an ongoing basis, we shared insights and actions from our human rights due diligence. Just prior to the reporting period, Meta disclosed a [detailed human rights report guiding the creation of the Oversight Board](#). During the reporting period, Meta disclosed due diligence on [Cambodia](#), [Indonesia](#), the [Philippines](#), [Sri Lanka](#), and — just after the reporting period — on its transition of Messenger and Instagram Direct Messages to [End-to-End Encryption](#).

05. For more on HRIAs, see Part 02.II of this report. Meta’s reference in this Report to third party diligence assessments cannot be construed as admission, agreement with, or acceptance of any of the findings, conclusions, opinions or viewpoints identified in those assessments, or the methodology that was employed to reach such findings, conclusions, opinions or viewpoints. Likewise, while Meta references steps it has taken, or plans to take, which may correlate to points assessors raised or recommendations they made, these also cannot be deemed an admission, agreement with, or acceptance of any findings, conclusions, opinions or viewpoints.

06. See also, UNGP 18 (“In order to gauge human rights risks, business enterprises should identify and assess any actual or potential adverse human rights impacts with which they may be involved either through their own activities or as a result of their business relationships. This process should: (a) Draw on internal and/or independent external human rights expertise; (b) Involve meaningful consultation with potentially affected groups and other relevant stakeholders, as appropriate to the size of the business enterprise and the nature and context of the operation.”); OHCHR, [The Corporate Responsibility to Respect Human Rights: An Interpretive Guide](#) (2012), at 33 (“Human rights due diligence is about people. It reflects the entitlement of every human being to be treated with dignity. It therefore involves relationships — between an enterprise and those on whom it may have an impact. Hence, the key to human rights due diligence is the need to understand the perspective of potentially affected individuals and groups. Where possible and appropriate to the enterprise’s size or human rights risk profile, this should involve direct consultation with those who may be affected or their legitimate representatives, as discussed further under Guiding Principle 18.”).

Meta's Salient Risks

Every day, billions of people all over the world use Meta's social media platforms and messaging apps.

In this report, we aim to show the most salient human rights issues identified by our due diligence efforts to date, the actions taken across product and policy teams to prevent or mitigate these risks, and how we developed processes to assess the effectiveness of those measures.



“Human rights risks are understood to be the business enterprise’s potential adverse human rights impacts.

Potential impacts should be addressed through prevention or mitigation, while actual impacts — those that have already occurred — should be a subject for remediation.”

COMMENTARY TO UNGP 17

A. How Meta Identifies its Salient Risks

As the Key Concepts section of the [Interpretive Guide to the UNGPs](#) clarifies:

The most salient human rights for a business enterprise are those that stand out as being most at risk. This will typically vary according to its sector and operating context. The Guiding Principles make clear that an enterprise should not focus exclusively on the most salient human rights issues and ignore others that might arise. But the most salient rights will logically be the ones on which it concentrates its primary efforts.

Meta uses a thorough human rights assessment process to identify salient risks in a given or particular context. Drawing on these assessments — as well as our ongoing day-to-day human rights due diligence — we established an initial list of salient human rights issues.

We commit to conducting a company-wide salient risk analysis in the near future and reporting on it in a forthcoming annual report.



B. Salient Risks

Based on due diligence efforts to date, the categories of salient risks for Meta can include, for example:

Rights	Human Rights Standards
Right to Freedom of Opinion and Expression	Universal Declaration Of Human Rights , Art. 19 International Covenant On Civil And Political Rights , Arts. 19, 20 Convention On The Rights Of The Child , Arts. 13, 17 International Convention On The Elimination Of All Forms Of Racial Discrimination , Art. 4
Right to Privacy	Universal Declaration of Human Rights , Art. 12 International Covenant on Civil and Political Rights , Art. 17
Right to Life, Liberty and Security of Person	Universal Declaration of Human Rights , Art. 3 International Covenant on Civil and Political Rights , Arts. 6, 9 Convention on the Rights of the Child , Art. 6
Rights to Equality and Non-Discrimination	Universal Declaration of Human Rights , Arts. 1, 2, 7 International Covenant on Civil and Political Rights , Arts. 2, 3, 26 International Covenant on Economic, Social and Cultural Rights , Arts. 2, 3 Convention on the Elimination of All Forms of Discrimination Against Women , Art. 2 International Convention on the Elimination of All Forms of Racial Discrimination , Art. 2
Best Interests of the Child	International Covenant on Civil and Political Rights , Art. 24 Convention on the Rights of the Child , Art. 3
Rights to Public Participation, to Vote, Be Elected	Universal Declaration of Human Rights , Art. 21 International Covenant on Civil and Political Rights , Arts. 25 Convention on the Elimination of All Forms of Discrimination Against Women , Art. 7, 8 International Convention on the Elimination of All Forms of Racial Discrimination , Art. 5

PART 2:

Meta's Human Rights Policy in Practice

Our approach to human rights is grounded in the five pillars of our [Human Rights Policy](#):

- I. Applying Relevant Policies;
- II. Conducting Human Rights Due Diligence and Disclosure;
- III. Providing Remedies for Human Rights Impacts;
- IV. Protecting Human Rights Defenders; and
- V. Governance, Oversight and Accountability

As we reflect on the Policy's first year, we address our progress and our room for growth across each area of commitment.

Applying Relevant Policies

Under our [Human Rights Policy](#), our respect for human rights — and their underlying principles of equality, safety, dignity, privacy and voice — is applied through our existing policies, including the Facebook [Community Standards](#), Instagram [Community Guidelines](#), [WhatsApp Terms of Service](#), the [Responsible Supply Chain](#) program, our [Privacy Principles](#), our [Law Enforcement Guidelines](#), our [Data Policy](#), our [WhatsApp Privacy Policy](#), our [Code of Conduct](#), and our [Diversity and Inclusion](#) practices, among other policies so included.⁷

In the following section, we highlight key human-rights related developments in the application of our policies.



7. Moreover, as discussed *infra*, we extend our human rights commitments to additional, internal product-focused decision-making approaches. These efforts require coordination, dependencies and alignment with Products, Operations, and Corporate partners.

A. Content Policies

Freedom of expression is a fundamental right, one that is interdependent with other fundamental rights. We look to international human rights experts when developing our standards for what content is and is not allowed on our social media platforms, and when deciding how to implement these standards in practice.

These rules are known as the [Community Standards](#) for Facebook and [Community Guidelines](#) for Instagram. They were — and

are — developed based on feedback from our community and the advice of experts in fields such as technology, public safety [and human rights](#). In seeking to ensure that everyone's voice is valued equally, we take care to create standards that include different views and beliefs, especially from people and communities that might otherwise be overlooked or marginalized. For these reasons, when we limit expression, we do it in service of one or more of a set of values: Authenticity, Safety, Privacy, and Dignity.

Our Commitment to Voice:

The goal of our Community Standards is to create a place for expression and give people a voice. Meta wants people to be able to talk openly about the issues that matter to them, even if some may disagree or find them objectionable.

THE FACEBOOK COMMUNITY STANDARDS

The full set of Community Standards are [available](#) in 66 language options to-date, as are our [detection and enforcement policies](#). Details on the effectiveness of our enforcement of these policies can be found in our quarterly [Community Standards Enforcement Reports](#).

Our work to infuse our Community Standards with human rights principles is ongoing and responsive to changing global salient risks, consistent with Meta's iterative approach to policy development.



Right to Freedom of Opinion and Expression

[Universal Declaration of Human Rights](#), Art. 19

[International Covenant on Civil and Political Rights](#), Arts. 19, 20

[Convention on the Rights of the Child](#), Arts. 13, 17

[International Convention on the Elimination of All Forms of Racial Discrimination](#), Art. 4

Our Community Standards and Community Guidelines are highly informed by international human rights standards (including the [Rabat principles](#)). Writing these policies involves extensive stakeholder engagement, including with human rights activists and academics. The implementation of these policies is informed by trusted partners who provide important context, especially in high-risk environments.

For example, Meta drew on international human rights law in determining how to address the complicated issue of [bullying and harassment of public figures](#). For this issue, the right to freedom of expression and legitimate public discourse around people in the public eye must be carefully balanced against public figures' rights to security and non-discrimination, honor, and reputation.

To develop a consistent approach, Meta consulted with a [diverse set of global stakeholders including](#) free speech advocates, human rights experts, women's safety groups, our Women's Safety Expert Advisors, cartoonists and satirists, female politicians and journalists, representatives of the LGBTQ+ community, content creators and public figures.

Based on stakeholder feedback, and guided by international human rights principles, the resultant [Community Standards Bullying and Harassment policy](#) maintains a distinction between public figures and non-public figures, and requires removal of [attacks on public figures](#) that encompass a wide range of harms to their fundamental rights. During the reporting period, this policy was further updated to remove certain severe sexualizing, degrading or derogatory content. The policy also provides additional protections to public figures like journalists and human rights defenders who have become famous involuntarily or because of their work.

In another, very important example: when a government sends Meta a take-down request, we typically conduct three levels of [review](#): against the Community Standards, against applicable local law, and against international human rights standards and our company [principles](#). No matter which government makes a given request, we will scrutinize it to make sure it is legally valid, proportionate and legitimate. In the interests of full transparency, we publish reports that set out statistics on [content we restricted based on local law](#), including in response to government requests. We also publish case studies of content we have restricted based on government takedown requests and [recently announced](#) we may also share these requests with [Lumen](#).



Right to Privacy

[Universal Declaration of Human Rights, Art. 12](#)

[International Covenant on Civil and Political Rights, Art. 17](#)

Privacy is one of the core values of the [Community Standards](#). As we note: it “gives people the freedom to be themselves, choose how and when to share on Facebook and connect more easily.” In addition to prohibiting personally identifiable information such as social security numbers, or other personal information directly identifying an individual, the Community Standards also prohibit the posting of personal contact information about oneself or others, including phone numbers or financial information about businesses or organizations, except when publicly available. In 2021, Meta requested a policy advisory opinion from the Oversight Board on the sharing of private residential addresses and images. In 2022, the Board recommended that we [disallow](#) this kind of sharing, and Meta [acceded](#).



Right to Life, Liberty and Security of Person

[Universal Declaration of Human Rights, Art. 3](#)

[International Covenant on Civil and Political Rights, Arts. 6, 9](#)

[Convention on the Rights of the Child, Art. 6](#)

We [expanded our policies that prohibit veiled and implicit threats](#). We added a policy to address hate speech attacks on concepts, ideas, practices, beliefs and institutions related to protected characteristics when those attacks pose an imminent risk of harm, intimidation or discrimination. We enhanced our bullying and harassment policy to include [stronger protections against gender-based harassment](#) for everyone, including public figures. We created new policies against certain [mass harassment or “brigading.”](#) We [also revisited our violence and incitement policies](#) — specifically, the allowance around the discussion of State use of force.



Rights to Equality and Non-Discrimination

[Universal Declaration of Human Rights, Arts. 1, 2, 7](#)

[International Covenant on Civil and Political Rights, Arts. 2,3, 26](#)

[International Covenant on Economic, Social and Cultural Rights, Arts. 2, 3](#)

[Convention on the Elimination of All Forms of Discrimination Against Women, Art. 2](#)

[International Convention on the Elimination of All Forms of Racial Discrimination, Art. 2](#)

We make country-specific decisions informed by [extensive stakeholder engagement](#) through our work with [Trusted Partners](#) and human rights activists and academics. Using existing policies and relationships, we improved our abilities to enforce at scale.

In 2021, we provided [an update on our work in Ethiopia](#) identifying and removing a number of persistent harmful false claims and out of context imagery that make false allegations about the perpetrators, severity or targets of violence in Ethiopia. This decision was based on guidance from over 50 local partners and independent experts.



Best Interests of the Child

[International Covenant on Civil and Political Rights](#), Art. 24
[Convention on the Rights of the Child](#), Art. 3

The [Philippines HRIA](#) spurred improvements to our existing policy language prohibiting multiple forms of human exploitation, and enforcement of our policy that [prohibits](#) all sales of people, including children. As we described in [Meta's Response](#), we updated policies around our proactive detection tools with the goal to reduce discoverability of exploitative content and incorporate emerging trends in consultation with the Philippines Department of Health and the National Kidney and Transplant Institute, as well as the Inter-Agency Council Against Trafficking, Overseas Workers Welfare Administration, Blas Ople Policy Center, Department of Foreign Affairs and the Philippine Overseas Employment Administration. We have ongoing coordination with the Philippines Inter-Agency Council Against Trafficking, Department of Justice, Department of ICT and the Department of Education as well as organizations advocating child safety, to quickly respond to reports and gather local insights and trends, and help us improve our policies, tools, and resources to better protect children online.



Rights to Public Participation, to Vote, Be Elected

[Universal Declaration of Human Rights](#), Art. 21
[International Covenant on Civil and Political Rights](#), Arts. 25
[Convention on the Elimination of All Forms of Discrimination Against Women](#), Art. 7, 8
[International Convention on the Elimination of All Forms of Racial Discrimination](#), Art. 5

Our [human rights-informed approach to election misinformation](#) balances freedom of expression against threats to the right to public participation. For example, [in preparation for the 2020 election in the United States](#), we took steps aimed at preventing the spread of misinformation, (including [clearer fact-checking labels](#)), fighting voter suppression and interference (including [banning paid ads that suggest voting is useless or advise people not to vote](#)), and helping people better understand the information they see online (including an [initial investment of \\$2 million](#) to support a media literacy project). We also [partnered with third-party fact-checkers on WhatsApp](#).

02. Community Standards in Focus: Dangerous Organizations

One area that has drawn significant attention from human rights stakeholders has been our [Dangerous Individuals and Organizations Policy](#) (“DOI” Policy), which addresses networks of people who “proclaim a violent mission or are engaged in violence.” We do not want Meta to be a platform for hate.

While terrorism and hate are global issues, there is currently no globally recognized and accepted definition of terrorist or hate organizations. So we developed [the definitions](#) in our [DOI Policy](#) to assist our decision-making on enforcing against such organizations:

- Our designations are divided into three tiers that indicate the level of content enforcement, with Tier 1 resulting in the most extensive enforcement, because we believe these entities have the most direct ties to offline harm.
- We designate Dangerous Individuals and Organizations under our policy after a rigorous process that takes into account both online and offline behavior. During this process, we work to identify an organization’s goals and if it has a track record of offline violence.
- As for **terrorist groups**, we designate individuals and organizations based on their behavior and whether they engage in, advocate, or lend substantial support to purposive and planned acts of violence with the intent to coerce, intimidate and/or influence a civilian population, government, or international organization in order to achieve a political, religious, or ideological aim.
- In addition to the definition of terrorist groups, and in recognition of the political complexity associated with the issue, we define as **violent non-state actors** those groups who primarily direct violence toward state or military actors (e.g. military targets).
- As for **hate groups**: We ban groups that proclaim a hateful and violent mission from having a presence on Facebook and Instagram and we remove content that represents, praises or supports them. To date, we identified a range of groups across the globe

as hate organizations because they engage in coordinated violence against others based on characteristics such as religion, race, ethnicity or national origin and we routinely evaluate groups and individuals to determine if they violate our policy.

- In addition, our policy covers Militarized Social Movements, Violence-Inducing Conspiracy Networks, and individuals and groups banned for promoting hatred.

Our process for assessing organizations for possible designation includes structured review of each case by a wide range of relevant teams, including policy, legal, and security. As the [DOI Policy](#) demonstrates, our definitions and thresholds are agnostic to region or ideology. Every designation goes through the process.

14

Strategic network
disruptions to remove

23

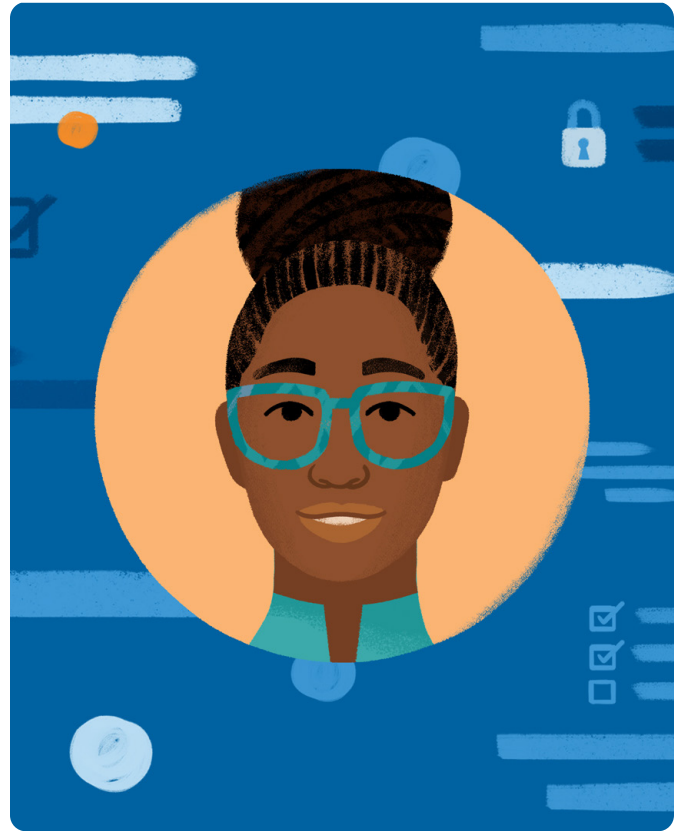
Different organizations
designated under our DOI policy

Finally, as part of our enforcement efforts against DOIs, we undertake certain strategic network disruptions, to help make sure that these groups cannot find ways back on our platforms and proliferate on our platforms. As outlined in our [Community Standards Reporting](#), between October 2019 and August 2020, we conducted 14 strategic network disruptions to remove 23 different organizations designated under our DOI Policy, including [three U.S.-based white supremacist groups](#) in October 2019.

B. Privacy

The right to privacy is a foundational right; an essential requirement for the realization of other rights, like the right to freedom of opinion and expression. Particularly in this digital age, interference with this right can both directly and indirectly limit the free development and exchange of ideas.

Privacy is one of the defining social issues of our time and is central to Meta's vision. We are [building a new privacy foundation to guide us now and in the future](#). We are committed to seeking feedback from and working with stakeholders across industry, civil society, think tanks and academia to improve our program.



Privacy Program Overview

Our [Privacy program](#) rests on a foundation of:



Governance: A cross-functional group of organizations across the company provides engineering, legal, policy, compliance and product expertise that enable the design and implementation of our privacy program. This includes an independent committee of Meta's Board of Directors that meets quarterly to oversee our efforts to live up to our privacy commitments.



Privacy Education. To make privacy a core responsibility of every Meta employee, we drive continuous privacy learning and education that spans training, internal campaigns, regularly provided privacy content and other dynamic resources. We do this through internal Workplace channels, lightning talks with privacy leadership, internal Q&A sessions, a dedicated Privacy Week and an internal hub of on-demand privacy content to help guide decisions and processes.



Accountability in Practice. In order to put our accountability foundation into practice, we designed processes, escalation paths and technical mechanisms that embed privacy across all facets of our company operations. This includes:

- **A privacy risk assessment program** that performs an annual assessment to identify, assess, and address privacy risk across the company, as well as a process to assess privacy risk after an incident occurs.
- **Safeguards and controls**, including operational activities, policies, and technical systems.
- A **privacy review process** by which we assess privacy risks that involve the collection, use, or sharing of people's information and external representations about our privacy and security practice; along with technical requirements and tools to enhance accountability and operate the privacy review process at scale, such as: a centralized tool that is used throughout the project lifecycle for Privacy Review and a technical implementation review that conducts review, verification, and documentation of the technical implementation of privacy mitigations and commitments prior to product launch.
- **Incident Management**, which includes: **proactive identification** of potential privacy issues, a **bug bounty** program that has, over the past 10 years, awarded bounties to around 1,500 researchers from 107 countries, and **transparency** efforts, like our Incident Management program, that includes steps to notify people where appropriate about issues impacting our community, or working with law enforcement or other officials to address what issues we find.
- **Third Party Oversight**, including a third-party privacy assessment process for service providers to assess and mitigate privacy risk at Meta, as well as a formal process for enforcing and offboarding third parties who violate our privacy or security obligations, such as procedures and infrastructure designed to ensure that third-party developers complete the Data Use Checkup, and other technical and procedural mechanisms to monitor compliance.
- **External Data Misuse.** We also have technical mechanisms in place to mitigate and prevent third parties from accessing data from Meta, through proactive and reactive measures like prevention, deterrence, detection and enforcement.



Privacy Product Outcomes. The accountability processes, safeguards, and technical mechanisms we built help ensure that new products and features embed privacy by design. We see these updated processes enable us to improve our privacy approach in new products and features, as we pivot to respond to the world around us. This includes:

- Giving people greater access and control over their information
- Giving people access to their information
- Giving people control to manage their activity
- Offering an ephemeral messaging option
- Creating a secure, private messaging experience
- Providing age-appropriate experiences for youth
- Communicating data privacy practices to people
- Using privacy enhancing technologies for social good



Investments In A Technical Foundation That Supports Privacy, *i.e.*, sustainable technical solutions to meet evolving privacy expectations and ensure consistent application of our privacy requirements across our products and systems, such as:

- A data deletion framework that helps alleviate the risk of potential error through machine-learning automation; and
- Privacy-enhancing technologies based on advanced cryptographic and statistical techniques that minimize the data we collect, process and share — to help protect data at different stages of the data lifecycle.

Our privacy work is never finished, and we understand that our commitment means continuously improving and focusing on this every day.





01. Salient Risks in Privacy

Two highlights of how Meta approaches its salient privacy risks include: protecting users from unlawful or overbroad government demand for user data; and our responsibility to protect user data and use it, share it, store it, and delete it pursuant to privacy and data protection principles and applicable legal requirements.

Meta's approach to both challenges builds on our theory of change: developing systemic, scalable, integrated analytical approaches that provide guardrails informing product development enterprise-wide.

The salient risk of government overreach and abuse in demanding user data from private technology companies was the starting point of Meta's human rights journey.

I. Protecting Users from Unlawful or Overbroad Government Data Requests

The salient risk of government overreach and abuse in demanding user data from private technology companies was the starting point of Meta's human rights journey. The salience of this risk has not diminished over time. In compliance with the GNI Principles, and our commitment to the UNGPs, we conducted human rights due diligence to develop an approach for our responses to government demands for user data.

As set out in Facebook's Data Policy, and guidelines for Law Enforcement, we scrutinize every government request we receive, no matter which government makes the request. There are many legitimate and lawful reasons why governments request information from social media companies: international criminal investigations, anti-terrorism work, prevention of child sexual abuse and other efforts intended to protect the lives, security and rights of people in their countries. We respond to such requests for information that are consistent with internationally recognized standards on human rights, including due process, and the rule of law. When we do

comply, we produce only the information that is narrowly tailored to that request. If we determine that a government request is not consistent with applicable law or our policies, or unlawful (for example, overly broad, or legally deficient in any way), we will push back.

We do not provide governments with "back doors" to people's information and we would challenge any order that sought to have us redesign our systems to undermine the encryption we provide to protect people's data, or any attempt to gag us from disclosing the existence of such an order and our efforts to fight it. For example, Facebook and WhatsApp paused — on human rights grounds — all data disclosure to the Hong Kong government after the passage of the National Security Law in 2020.

We "show, not tell," by regularly publishing transparency reports that set out, by country, the number of government requests we have received and the proportion of requests for which some data was produced.

Reform Government Surveillance Coalition

Along with other major tech companies, Meta is a member of the [Reform Government Surveillance Coalition](#), which urges the world's governments to adopt surveillance laws and practices consistent with established norms of privacy and free expression and the rule of law.

The [six core principles](#) of the coalition are:



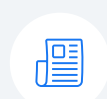
1. Limiting governments' authority to collect users' information;



2. Oversight and accountability;



3. Transparency about government demands;



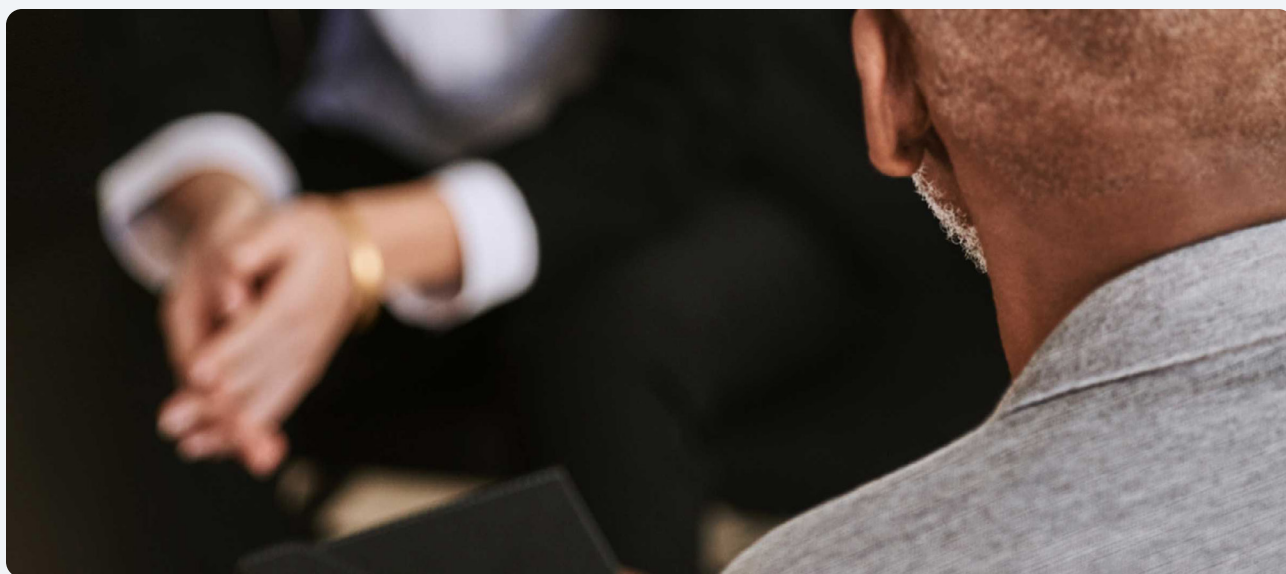
4. Respecting the free flow of information;



5. Avoiding conflicts among governments; and



6. Ensuring security and privacy through strong encryption.





II. Properly Collecting, Using, Storing, and Deleting User Data

We [made progress on our work to give people more control over their data](#), and our broader mission to honor people's privacy in everything we do by building processes, products, and technical mechanisms that laid the foundation for privacy and accountability across the company. The program includes internal education, due diligence, safeguards and controls, and third-party oversight. For example, we provide people with ways to:

- [Manage their information](#). We make it easy for people to access, manage, download or delete the personal data they have provided.
- [Transfer their data](#). We make it simple to move information, such as posts, photos and videos, to another service.
- [Access the Accounts Center](#). People can control the experiences they connect across our apps by using our Accounts Center.
- [Access the Instagram Activity Center](#). People can access and manage content they've commented on, shared or sent on Instagram.
- [Perform a Privacy Checkup](#). This tool helps people control who can see what they share,

how their information is used and how to secure their account.

- [See "Why Am I Seeing This Ad?"](#) This tool lets people tap on posts from the friends, Pages, and Groups they follow as well as some of the posts Meta suggests to get more context on why they are appearing in News Feed.
- [Review Off-Facebook Activity](#). We created a tool so people can control — or disconnect — the information businesses send to Meta about their activity on other apps and websites.
- [Review Ads Interests & Preferences](#). This tool gives people more visibility into the actions they've taken to control the topics of the ads you see.
- [Access Cookie Controls](#). In the European Region, we give people a more granular level of control over their cookie choices and more information on how we use different kinds of cookies.
- [Set up Two-Factor Authentication](#). We give people this option in addition to their password to help protect their account from improper access.

C. Civil Rights

Civil rights are human rights, but they are also understood to be personal rights protected by the Constitution of the United States and Federal and State laws. The Meta Civil Rights team focuses on issues of discrimination and exclusion based upon race, ethnicity, religion, national origin, gender, gender identity, sexual orientation and disability.

Meta's commitment to human rights and civil rights are closely interrelated. Our Human Rights Policy lays a strong basis for continuing civil rights and racial justice [work](#) and [standards](#).

The initial framework for our Civil Rights work was the [civil rights audit](#) we commissioned at the behest of the civil rights community to identify instances where the company could take a more comprehensive approach to respecting civil rights. Meta published the findings in full: a total of three reports, the [last of which](#) was published in July 2020. The extensive audit resulted in 117 recommendations and actions with input from over 100 civil rights and social justice organizations.

This audit laid the groundwork for the Civil Rights Team that was created in November 2020 and provided the company with an initial roadmap.

In November of 2021, Meta published a report that detailed [Meta's Progress on Civil Rights Audit Commitments](#), demonstrating success in implementing 65 of the Auditors' recommendations and actions, as well as demonstrating ongoing or in-progress status of 42, and continuing evaluation of the feasibility of eight. This Progress Report explained that Meta's civil rights work focuses on five key pillars:



Law Enforcement & Hate: Addressing harm and creating accountability are critical to enhancing protections for marginalized communities and victims of hate speech, hate incidents and hate crimes are safe on our platforms. It also means working to prevent law enforcement misconduct and surveillance of marginalized communities on our platforms as these activities can violate our policies and chill free expression rights.



Voting & Civic Engagement: Grounding the work in efforts to inform and connect people, communities, and civic institutions to come together and drive real outcomes to improve lives, because when all people have an opportunity to participate and are considered political equals, societies make decisions optimized for the many, not the few.



Technology: It is Meta's responsibility to help instill civil rights best practices throughout the company with regard to artificial intelligence, machine learning, virtual and augmented reality, and emerging technologies, and lead engagement with external technology experts.



Product: Meta aims to advance our priority of doing good and causing no harm by creating product review approaches and guidance tools aimed at prioritizing protected classes and protecting systemically and historically marginalized communities.



Policy & Enforcement: Meta works to surface civil rights concerns on the front end of policy development and enforcement decisions and ensure a more diverse and inclusive engagement process.

Growing out of the civil rights audit, the Meta Civil Rights Team has grown to include people with expertise spanning hate crimes, voting rights, counterterrorism, immigration, national security, law enforcement, product inclusion, algorithmic fairness and transparency.

As this team continues to advance civil rights and liberties, it does so in partnership with the Human Rights Policy Team. Meta is committed to continuing to build with the values of justice, equity, dignity, and safety in mind.



D. Responsible Innovation and Responsible AI

To help ensure that we build the future of the internet responsibly, Meta developed human rights-based integrated analytical approaches to support our innovation and product development.

Our [Responsible Innovation Dimensions](#) support product development. This framework is evolving over time, but currently includes 10 dimensions: autonomy, civic engagement, constructive discourse, economic security, environmental sustainability, fairness and inclusion, privacy and data protection, safety, voice, and well-being. These dimensions in turn guide our analysis and practice.

Artificial Intelligence (AI) is a core component of the technologies Meta uses to provide value to people and [keep our communities safe](#), from helping [rank posts in News Feed](#) to [tackling hate speech](#) and [misinformation related to the](#)

[COVID-19 pandemic](#). Meta also developed a [dedicated, cross-disciplinary Responsible AI \(RAI\) team](#) within its AI organization to help ensure that AI governance is based on foundational values of respect for human rights, democracy, and the rule of law.

Those foundational values are at the root of the wide range of principles statements that have been released around responsible AI development, most especially the European Commission's [High-Level Expert Group's Ethics Guidelines for Trustworthy AI](#) and the [Organization for Economic Cooperation and Development's Principles on Artificial Intelligence](#), which Meta helped develop.

Meta, in turn, organized its Responsible AI efforts around [five key pillars](#):



Privacy & Security



Fairness & Inclusion



Robustness & Safety



Transparency & Control



Accountability & Governance



We also invest in research and open-sourcing datasets and tools to help facilitate responsible use of AI, such as [privacy-preserving machine learning](#), [AI explainability](#), and [fairness](#). For example, in 2021 we released our [Casual Conversations data set](#), composed of over 45,000 videos designed to similarly help researchers evaluate computer vision and audio models for accuracy across a diverse set of ages, genders, apparent skin tones, and ambient lighting conditions.

We are also improving transparency by piloting simple, standardized documentation of our models and [using interpretability software such as Captum](#). Although work in this area is still in its infancy, [our hope](#) is that ultimately we will be able to build an integrated transparency solution that can automatically feed information from internal documentation efforts — like model cards⁸ — into new transparency features and controls for the people using our products.

In addition to our technical research and product-focused work, we are actively participating in efforts to establish clear AI principles and best practices, including [collaborating with the OECD on the AI Observatory project](#) to study and

disseminate emerging best practices that are in line with its [AI Principles](#).

Through our [partnership with Open Loop](#), we are building innovative “policy prototyping” projects for testing new potential AI policy requirements with regulators and startups before they become law, to ensure that they are both practical and impactful. [We launched](#) projects in Europe and the Asia Pacific region.

Finally, we are [funding a global effort](#) to solicit diverse academic research on AI ethics and governance topics, supporting the publication of academic papers in [Asia](#), [Africa](#), and [Latin America](#) and providing foundational support for an [independent Institute for Ethics in Artificial Intelligence at the Technical University of Munich](#).

08. Model Cards are short documents accompanying trained machine learning models that provide benchmarked evaluation in a variety of conditions, such as across different cultural, demographic, or phenotypic groups (e.g., race, geographic location, sex, Fitzpatrick skin type) and intersectional groups (e.g., age and race, or sex and Fitzpatrick skin type) that are relevant to the intended application domains. Model cards also disclose the context in which models are intended to be used, details of the performance evaluation procedures, and other relevant information.

E. Accessibility



The [UN Convention on the Rights of People with Disabilities](#) sets out eight fundamental principles that underpin the protection of the rights of people with disabilities. These include non-discrimination, full and effective participation and inclusion in society and accessibility. Given the importance of the digital world for people with disabilities to be able to participate fully in exercising their rights, Meta sought to apply cutting-edge technologies and make other investments to create more accessible features and services for users with a wide range of disabilities. Such measures include:

- [Applying the latest AI techniques to increase object and concept recognition in Automatic Alt Text](#) or “AAT,” our award-winning automatic photo-description technology that works with screen readers to describe photos to people who are blind or have low vision, by more than 10X.
- Re-engineering facebook.com to [integrate greater accessibility into the website](#), including technology that helps people using screen readers better understand what’s on a page and find what interests them more quickly.
- [WhatsApp has various accessibility features](#), including a voice-assisted camera feature, which makes it easier for people who are blind or have low vision to take photos on iOS devices, and a new WhatsApp email accessibility support channel for users to share suggestions, ask questions, or report problems.
- We introduced new accessibility-specific features in Quest 2, including color filters to assist people with color-sensitivity and “adjust height” to enable better VR experiences for people who are reclining, seated, or lying down.
- Automatically generated captioning available for [Facebook Live](#), [Workplace Live](#), [Instagram TV](#), and [Live Audio Rooms](#), as well as for Facebook ads, Pages, and Groups.

In addition, we engaged in several initiatives to help promote accessibility beyond Meta. We work collaboratively with leaders in the accessibility community globally, and are working within our industry and through the [W3C](#) to establish an interoperability standard for web and assistive technology.

We also co-founded and remain active in [Teach Access](#), a collaboration of industry, academia, and disability advocates to advance the teaching of accessibility in education. Students learn state-of-the-art best practices in accessible software design, preparing them to make future technologies that are “born accessible.” We are investing early in making the metaverse accessible, including by creating XR Accessibility guidelines for developers in partnership with the XR Association and additional guidelines specifically for third-party developers of the Quest VR platform.

F. Anti-Slavery and Human Trafficking

Meta is opposed to all forms of human trafficking, slavery, servitude, forced or compulsory labor, and all other trafficking-related activities. [We are committed](#) to respecting all applicable international human rights standards, labor and employment laws, rules, and regulations, and to working to mitigate the risks of modern slavery and human trafficking in our business operations and supply chains. We manage labor, human rights, and environmental risks in our supply chain through our [Responsible Supply Chain](#) (RSC) program. We collaborate and share ideas with the broader industry and global community to update and develop solutions that promote best practices for a responsible supply chain, and to that end we are members of the [Responsible Business Alliance](#) and the [Responsible Labor Initiative](#).

We require all personnel be trained on and comply with our [Code of Conduct](#), which affirms our commitments to human rights and prohibits violations of law, including labor and employment laws. Meta requires that priority suppliers, which are identified using a risk-based methodology, conform with all of the standards on labor, health and safety, the environment, business ethics, and the establishment of management systems

in accordance with the [Responsible Business Alliance \(RBA\) Code of Conduct](#), which includes standards related to indicators of modern slavery and human trafficking.

We [verify supplier conformance](#) with RSC policies and standards through continuous dialogue, independent audits and assessments, corrective action plans, worker surveys, and other forms of assurance. Any non-conformances identified are actively addressed through our corrective action and key performance indicator programs. We regularly assess and track the effectiveness of our actions taken in mitigating the associated risks of modern slavery and human trafficking. We strive to improve year-on-year; every year we publish a refreshed/new [statement](#) to demonstrate progress and strengthening of our program efforts.

As noted, Meta [strongly opposes the abuse of our family of apps and platforms](#) to facilitate any form of human exploitation. At the core of our strategy, we continually update and refine our policies to ensure we are keeping up with salient human rights risks related to how abuse tactics change over time.

Strategic partners include:

[National Center for Missing and Exploited Children](#), [International Center for Missing and Exploited Children](#), [Internet Watch Foundation](#), [ECPAT](#), [Polaris](#), [Stop the Traffik](#), and Meta's Expert Advisory Group

We implemented new in-product features to raise awareness of trafficking, deter violating behavior, and offer support to victims. One example is our search redirection that detects when people search for potentially violating keywords and delivers a custom deterrence message. We are increasing our investments and our ability to identify the illicit actors, networks, organizations, and businesses that perpetrate these activities and disrupt them accordingly.

We collaborate with global NGOs to create educational campaigns, provide resources and assist victims and survivors of human trafficking. We organize internal training, and events to raise awareness of human exploitation amongst our employees.

We continually seek to expand and refine our understanding of emerging salient risks related to human exploitation as we amend and update our enforcement guidelines.

G. Election Integrity

We are [committed to protecting elections](#), to increasing authenticity, transparency and accountability for advertisers, to giving everyone a voice on our technologies and to empowering people to vote. Between 2020 and 2021, Meta made significant investments in teams and technologies to better protect elections, empower people to vote and create safe technologies to share diverse beliefs.



01. United States Presidential Election 2020

For the United States Presidential Election in 2020, we started preparations two years in advance. These enabled us to identify emerging threats and put systems in place to mitigate anticipated risks.

Among the threats we expected to see in the lead-up to the November election were: perception hacking, the shift from larger-scale operations to narrower campaigns that try to slip under the radar and leverage unwitting authentic people, and also blurring lines between authentic public discourse and manipulation by co-opting domestic groups to amplify and join influence operations. [Actions taken included:](#)

- We removed almost a dozen foreign operations backed by Russia, China and Iran that used fake accounts to deceive users and undermine trust in the United States under our policy against coordinated inauthentic behavior, announced each of these takedowns publicly, and shared information with independent researchers so that they could review and draw independent findings.
- We removed more than 265,000 pieces of content on Facebook and Instagram in the US between March 1 and Election Day for violating our voter interference policies.
- We attached informational labels to content that discussed the legitimacy of the election or claimed that lawful methods of voting like mail-in ballots would lead to fraud.
- We ran the largest voting information campaign in American history, connecting people with reliable information about voting from state and local election authorities as well as nonpartisan civic partners.
- We [worked](#) with law enforcement in the days and weeks after January 6, 2021 and continue to do so with the goal of ensuring that information linking the people responsible for their crimes is available.

02. Myanmar Election 2020

Having identified Myanmar as high priority and high risk in advance of its November 2020 election, Meta created a multi-disciplinary planning team with a human rights presence to help identify, prevent or mitigate key salient human rights risks. Meta also [took steps to prepare](#) for the election, including:

- Stood up a dedicated Integrity Product Operations Center with 24/7 coverage during the election to respond to risks and on/offline developments in real-time.
- Expanded our misinformation policy in Myanmar in an effort to help combat voter suppression, and help support the integrity of the electoral process, including working with local partners to remove verifiable misinformation and unverifiable rumors (e.g., posts falsely claiming a candidate is a Bengali, not a Myanmar citizen, and thus ineligible).
- Required all electoral and political ads in Myanmar have a “Paid for by” disclaimer attached to them to show the organization or person behind the ad.
- Worked with partners in Myanmar to verify official Facebook Pages of political parties and candidates.
- Worked with civil society partners to surface and rapidly review potentially harmful content, including misinformation that could lead to offline harm, veiled threats, and hate speech.
- Introduced an Image Context reshare product which provided people using the platform in Myanmar with additional context before they shared images that were more than a year old and could be potentially harmful or misleading.
- Introduced a new feature that limits the number of times a message can be forwarded to five. This safety feature was made available on Messenger in Myanmar in 2020.
- Utilized hate speech classifiers to proactively detect probable hate speech in Burmese and either demote, route for human review, or automatically delete depending on confidence level.
- Implemented proactive detection to surface content containing certain keywords/phrases associated with hate, misinformation, and other harms for further review.
- Introduced a third-party fact-checking program in Myanmar as part of our ongoing integrity efforts to reduce the spread of misinformation and improve the quality of the news people find online.
- Directed people to authoritative sources of election information, where they could learn how to check voter lists, as well as voting times and locations, including Union Election Commission, Vote MM and First Time Youth Voters for 2020.
- Trained civil society organizations and reporters on journalist safety, media and digital literacy, as well as Meta’s policies and third-party fact-checking programs. As part of this effort, we held a monthly television talk-show on digital literacy called Tea Talks, that focused on issues like online bullying and account security.
- Introduced tools to newsrooms in Myanmar such as CrowdTangle, a public insights tool that makes it easy to follow, analyze and report on what’s happening with public content on social media.
- Identified and disrupted six networks engaging in Coordinated Inauthentic Behavior in Myanmar. These networks of accounts, Pages and Groups were masking their identities to mislead people about who they were and what they were doing by manipulating public discourse and misleading people about the origins of content.

03. Preparation for Philippines Presidential Election 2022

In line with our approach to other major global elections, Meta made it a high priority to invest in risk mitigation and preparation efforts for the 2022 presidential election in the Philippines. These efforts were informed and guided by recommendations of the [Philippines Human Rights Impact Assessment](#), and were a demonstration of our [commitment to follow up](#) on that work. We built new products and developed stronger policies in collaboration with the Commission on Elections, election watchdogs, independent fact checkers and civil society organizations. [In addition](#):

- In partnership with the Philippine Commission on Elections and various civic organizations including the Legal Network for Truthful Elections (LENTE), we launched a civic education campaign, conducted training on Meta's policies, and held roundtables with civil society.
- We've expanded our flagship digital literacy program, Digital Tayo to reach over 6.5 million people in the Philippines. Digital Tayo covers topics such as online safety, privacy, digital citizenship, news and media literacy, and launching civic campaigns.
- We supported Internews Philippines to launch the Philippine Fact Checker Incubator, to increase the capacity of local organizations for fact checking.
- To help people better understand who is behind the news they see on Facebook and Instagram, we applied labels to media outlets that are determined to be wholly or partially under the editorial control of their government, including those in the Philippines.
- As part of tackling other emerging harms, we also removed a network of over 400 accounts, Pages, and Groups in the Philippines that worked together to systematically violate our Community Standards and evade enforcement.



H. COVID-19

The world has suffered immense health and human rights consequences since COVID-19 was [declared by the WHO](#) to be a Public Health Emergency of International Concern (PHEIC) in January of 2020.

Throughout the reporting period, Meta mobilized to support public health, amplify authoritative information, connect users to essential services, as well as assist relevant agencies in their life-saving work. Three areas of Meta's COVID-19 response are particularly relevant to this report: (1) emergency due diligence, (2) content policy due diligence, and (3) additional actions to support the right to health.⁹

In 2021, we also produced a [detailed white paper](#) summarizing actions Meta took to support effective COVID-19 response while protecting privacy.

09. Further details of Meta's COVID response are available at our [Responding to COVID 19](#) site, and on the [newsroom](#)



01. Emergency Due Diligence

As the pandemic broke out, the Human Rights Policy Team worked on an emergency footing with other internal teams to align emergency response actions with UNGP guidance and prioritization frameworks. Their [actions included](#): real time decision support and advice related to the rights to health, information, and freedom of

expression; as well as rapid review of emergency product measures, adjustments to automated systems, adjustments to the at-risk countries workstream, and frameworks for prioritizing content for human review by a temporarily reduced pool of content moderators.

02. Content Policy Due Diligence

As the pandemic spread, Meta's content policy teams worked to respect and incorporate authoritative guidance related to the right to health, the right to information, and [permissible limits on freedom of expression](#).

Informed by this and other work, Meta took aggressive steps to limit the spread of misinformation about the virus and its effective vaccines and treatments, as well as to connect people with reliable health and service information. We also updated other relevant content policies to mitigate COVID-19-related risks, including (among others) our hate speech policies; coordinating harm policies; restricted goods policies; and bullying and harassment policies.

The full set of our relevant policies are available in: the [Misinformation section of the Facebook Community Standards](#), the [COVID-19 and Vaccine Policy Updates & Protections section of the Facebook Help Center](#), by the [Instagram Community Guidelines](#), and the Instagram-specific page relevant to [COVID-19 and Vaccine Policy Updates and Protections](#).

Under these policies, we remove content containing claims related to COVID-19 and vaccines or treatments that, according to public health authorities, are (a) false, and (b) likely to contribute to imminent physical harm. Imminent physical harm examples include: increasing the likelihood of exposure to or transmission of the virus, or having adverse effects on the public health system's ability to cope with the pandemic.

Based on input from experts in health communication and related fields, we took additional steps to reduce the distribution of content that does not violate Meta policies but may present misleading or sensationalized information about vaccines in a way that would be likely to discourage vaccinations.

In an important step for enforcement at scale, Meta [developed and deployed](#) AI tools to scale fact-checker's work and detect copies of rated content when someone tries to share them. We also built new computer vision classifiers to enforce bans on violating ads and commerce listings for certain medical products.



of the time people who saw the warning label did not click past to view

These efforts were effective. In April 2020, for example, [we put warning labels on about 50 million pieces of content based on around 7,500 fact-checks from partners](#). During the month of March 2020, we displayed warnings on about 40 million posts related to COVID-19 on Facebook, based on around 4,000 articles by our

independent fact-checking partners. 95% of the time people who saw the label did not click past to view. From the beginning of the pandemic to July 2021, we put these warning labels on [167 million pieces](#) of content.

We also made significant efforts to serve public health organizations and emergency responder organizations, globally and locally. We connected over 2 billion people from 189 countries to

resources from health authorities through our COVID-19 Information Center and pop-ups on Facebook and Instagram. We provided [ad credits](#) to governments and health agencies while WhatsApp enabled national health ministries around the world to communicate at scale with their citizens by onboarding them to the WhatsApp Business API, allowing them to provide timely information and respond automatically to frequently asked questions.

03. Privacy Due Diligence and Data for Good

Our [Data for Good](#) Program has been sharing privacy-protective tools and data with its 600+ partners in more than 75 countries, many of whom are working to track the spread of the virus; measure the effectiveness of prevention measures; and better understand attitudes towards vaccines.

Meta sought to address the urgent needs of public health authorities while maintaining a strong commitment to privacy and other fundamental rights. After committing to be guided by public health experts in our response, Meta decided to prioritize areas where our existing strengths made our contributions particularly impactful.

Meta understood from epidemiologists that surveys, in particular, could assist their efforts in three ways: (1) forecasting the spread of the virus; (2) assessing peoples' understanding of preventing behaviors, such as mask wearing; and (3) understanding attitudes and behaviors related to vaccinations as vaccines became available.

As a result, we launched a number of mobility datasets that were used across dozens of countries to better understand the effectiveness of lockdown orders and partnered with select academic and public health partners to create a

large scale COVID-19 Trends and Impact Survey, to which over 100 million people in 130 countries and territories responded. We also hosted public visualizations of this publicly available information to help communicate to a wide range of audiences.

While Meta was developing its strengths-based approach, the human rights and privacy teams undertook detailed human rights due diligence to guide Meta positions on responding to governmental COVID-relevant data sharing requests, the development and promotion of contact-tracing interventions, and the collection of sensitive data via chatbots in encrypted messaging services. These complex frameworks confirmed the importance of the strength-based approach. Meta did not support the promotion of automated contact tracing apps and did not seek to develop such technology, focusing its efforts instead on sharing population-level insights on changes in mobility and how this might affect transmission and the effectiveness of stay at home orders around the world.

More detailed information on these initiatives, along with key discussion questions, is available in our April 2021 white paper: [A Retrospective — Protecting Privacy in Our Covid-19 Response](#).

Conducting Human Rights Due Diligence and Disclosure

Due diligence is the second core pillar of Meta's Human Rights Policy. Pursuant to the UNGPs, we endeavor to (a) assess actual and potential human rights impacts; (b) take appropriate action and implement findings; (c) monitor implementation; and (d) communicate how impacts are being addressed. We seek to identify vulnerable or marginalized groups and engage, meaningfully, to listen to concerns, and create solutions. As part of our "show, not tell" ethos, we have sought to demonstrate meaningful transparency in our due diligence exercises.

In this section, we detail examples of the innovative approaches to diligence that we have operationalized, as described earlier:

1. Human rights impact assessments of specific countries, product initiatives, or strategic initiatives;
2. Integrated analytical approaches embedding UNGP prioritization criteria; and
3. Rapid salience analyses to support crisis situations, real-time content policy feedback, and product decision-making.

As the UNGPs were only approved by the UN Human Rights Council in 2011, the field of human rights due diligence is nascent, with rapidly evolving methodologies and best practices. For social media companies like Meta, whose products, scale of operations and impact are also evolving at a dramatic pace, the challenge of iterating methodologies capable of identifying salient risks and developing meaningful recommendations is considerable.





A. Human Rights Impact Assessments

A human rights impact assessment (HRIA) is a detailed, direct form of human rights due diligence that allows companies like Meta to identify potential human rights risks and impacts, to promote human rights and seek to prevent and mitigate risks. This work is in line with both our commitments as members of the GNI, and our responsibility under UNGP 18.¹⁰

In 2018, Meta commissioned and published the full text of an independent HRIA for Myanmar: [An Independent Assessment of the Human Rights Impact of Facebook in Myanmar](#).

In May 2020, we [published executive summaries of and recommendations from three more HRIAs for: Indonesia, Sri Lanka and Cambodia](#), along with our responses ([Indonesia](#), [Sri Lanka](#), [Cambodia](#)), at which time we also directly briefed interested and affected rights-holders, with the exception of Cambodia where COVID-19 and the security environment prevented us from doing so.

Rights-holders, in this context, included not only users of our platforms and services, but the many others whose rights were potentially impacted by online activity and conduct.

Countries targeted for human rights due diligence exercises have stemmed from Oversight Board recommendations, recommendations from prior diligence exercises, and other ad-hoc determinations based on stakeholder feedback. We are moving to a more systematic approach informed by decision frameworks and stakeholder engagement.

10. UNGP 18 requires businesses, “in order to gauge human rights risks” to “identify and assess any actual or potential adverse human rights impacts with which they may be involved either through their own activities or as a result of their business relationships” while drawing on “internal or external expertise and “meaningful consultation with potentially affected groups and other stakeholders.”

01. Philippines HRIA



Salient Risks:

The [Philippines HRIA](#) was conducted by an independent third party from February to July 2020 and published in December 2021. The assessment found that Meta technologies are widely used and have positive human rights impacts in the Philippines, as well as salient human rights risks. For example, the assessment found that Meta technologies play an important and positive role in providing access to

economic opportunities, giving voice to people and being essential tools for monitoring and defending human rights during the COVID-19 pandemic. It also highlighted salient human rights risks, including concerns about the misuse of our technologies for misinformation and disinformation, online harassment, incitement of violence, surveillance, online sexual exploitation, human organs trafficking, and extremist activities.

Recommendations and Meta's Response:

The HRIA made 40 recommendations. At the end of 2021, Meta [committed to implement 24 recommendations](#), partly implement seven recommendations, and was assessing the feasibility of another nine. We will be tracking our implementation regularly through 2022 and beyond. In response to the HRIA recommendations,

we made substantial commitments to prepare for the 2022 Philippine presidential election and strengthened our partnerships and programs with Philippine civil society organizations and government entities. Meta's full commitments are available at [Meta Response: Philippines Human Rights Impact Assessment](#).

02. Summary Disclosure - India Human Rights Impact Assessment¹¹



Purpose

In late 2019, Meta commissioned an independent human rights impact assessment (HRIA) on potential human rights risks in India related to its platforms. The project was undertaken by Foley Hoag LLP¹². The project launched in March 2020. It experienced limitations caused by COVID-19, with a research and content end date of June 30, 2021.

The assessment was conducted independently of Meta. Meta provides this summary of the report's insights and actions consistent with commitments made in its Corporate Human Rights Policy, and the United Nations Guiding Principles on Business and Human Rights, including Principle 21(c).

This summary cannot be construed to endorse, accept or adopt the independent assessor's conclusions, recommendations or analysis.¹³

11. This summary was developed after consulting The Danish Institute for Human Rights Guidance on HRIA of Digital Activities Reporting and Evaluation

12. Foley Hoag LLP is a US law firm with a human rights practice.

13. Meta's publication of this summary, and its response thereto, cannot be construed as admission, agreement with, or acceptance of any of the findings, conclusions, opinions or viewpoints identified by Foley Hoag, or the methodology that was

employed to reach such findings, conclusions, opinions or viewpoints. Likewise, while Meta in its response references steps it has taken, or plans to take, which may correlate to points Foley Hoag raised or recommendations it made, these also cannot be deemed an admission, agreement with, or acceptance of any findings, conclusions, opinions or viewpoints.



Methodology

The HRIA involved interviews with 40 civil society stakeholders, academics, and journalists. Stakeholder identities were kept anonymous and not shared with Facebook/Meta. The due diligence also involved the review of content policies, certain relevant content, as well as a survey, designed across multiple dimensions of diversity, of over 2000 individual rights holders.

The report refers extensively to international standards. These include the United Nations Guiding Principles on Business and Human Rights; the U.N. International Bill of Human Rights (including the Universal Declaration of Human Rights (“UDHR”), the International Covenant on Civil & Political Rights (“ICCPR”), and the International Covenant on Economic, Social and Cultural Rights (“ICESCR”); the U.N. Rabat Plan of Action, including freedom of expression, information, and opinion; the right to non-discrimination; and standards relating to advocacy of national, racial, or religious hatred.



Limitations

The project experienced certain limitations caused by COVID-19. HRIAs conducted on the operations and activities of digital products exist in an emerging and constantly adapting space.

Based on the guidance of human rights experts, we have produced this synthesis in order to mitigate security risks as per UNGP 21(c).

While commissioned by Meta, the independent Assessment is not adopted by Meta or a statement on its behalf. Instead, the Assessment will ideally serve as a baseline to help identify and allow Meta to address the most salient platform-related human rights risks, and help guide additional human rights due diligence.

This document is intended to be a summary of the Assessment.



Context

In 2019, civil society groups published several reports criticizing Facebook content policy rules and content moderation processes in India. Using the guidance of the UNGPs, Meta initiated a human rights due diligence project to identify and mitigate potential human rights risks.



Insights

The HRIA was an independent report by a third party.

It found Meta platforms had provided an invaluable space for civil society to organize and gain momentum, provided users with essential information and facts on voting, and also enabled important public health updates. It also noted that Meta had committed to platform safety and integrity in India, and had devoted considerable energy in recent years to promoting respect for its users, and to protecting user safety.

These efforts included: implementing an industry-leading set of content rules, developed in close consultation with the human rights community; substantial investments in new resources to detect and mitigate hateful and discriminatory speech; new content moderation resources, and recruitment of expert global human rights personnel.



Risks

The HRIA noted the potential for Meta's platforms to be connected to salient human rights risks caused by third parties, including: restrictions of freedom of expression and information; third party advocacy of hatred that incites hostility, discrimination, or violence; rights to non-discrimination; as well as violations of rights to privacy and security of person. It stated Meta faced criticism and potential reputational risks related to risks of hateful or discriminatory speech by end users.

The Assessment also noted a difference between company and external stakeholder understandings of content policies. It noted persistent challenges relating to user education; difficulties of reporting and reviewing content; and challenges of enforcing content policies across different languages. In addition, the assessors noted that civil society stakeholders raised several allegations of bias in content moderation. The assessors did not assess or reach conclusions about whether such bias existed.



Actions

The HRIA developed recommendations covering implementation and oversight; content moderation; and product interventions; and other areas.

As previously stated, the conclusions and recommendations made in the Assessment are not statements made on behalf of or adopted by Meta. Meta is studying these findings, however, and will consider them as a baseline to identify and guide related actions. We continue to work towards providing a safe and open platform to all users.

On implementation and oversight, Meta has taken steps to expand its human rights team; aims to ensure there is appropriate regional expertise; and will seek to build awareness of its diverse slate recruitment approach globally. We'll deepen our stakeholder engagement in a comprehensive manner.

On content moderation, Meta already actively seeks to align our content moderation and enforcement policies with international human rights law.

In 2021—22, we made content policy changes related to a) groups at risk of outing under our [coordinating harm](#) policies and b) context-specific content attacking concepts and practices associated with protected characteristics. We will also develop a prototype test for application of the Rabat Principles, and will implement it for use as a policy tool of last resort.

In addition, in 2020—2022 Meta significantly increased its India-related content moderation workforce and language support. As of the time of writing, Meta had reviewers across 21 Indian languages, with both language and cultural expertise, as well as multiple classifiers. Translating widely used languages has substantial practical benefit for our users.

With regard to transparency on content trends, Meta began publishing detailed India transparency reporting in June 2021. The [India monthly report](#) is available at our [Transparency Center](#) and covers multiple policy areas for Facebook, Instagram, and [WhatsApp](#). In addition, we are actively seeking ways to improve our transparency reporting around government requests to remove or restrict content, as reported to the Oversight Board.

We will expand participation in our [Resiliency Initiative](#), which empowers local communities with digital tools to combat hate, violence, and conflict within and beyond their networks. We will also seek to expand our counterspeech [Search Redirect Program](#), which redirects users who search using hate and violence-related search terms towards resources, education, and outreach groups that can help.

With regard to product Interventions and other measures:

In 2021—2022, Meta designed and deployed multiple product features to try to prevent or mitigate acts of abuse on WhatsApp. Meta also disclosed a detailed independent human rights impact assessment of its expansion of End-to-End Encryption in April 2022. The recommendations will help guide our [approach to safer private messaging](#): helping to prevent abuse and to safeguard people's privacy, and giving people controls to help them stay safe.

We are developing consultation processes around our annual human rights disclosures at the time of writing. We are implementing mergers and acquisitions human rights due diligence, to be incorporated into our standardized due diligence processes. We are launching our human rights website in July 2022. As per our corporate Human Rights Policy, we are sharing insights and actions from this due diligence in our annual human rights report.

B. Product-Focused Human Rights Impact Assessments

Meta commissions detailed human rights impact assessments in instances of ground-breaking new products, or systemic product changes to

current platforms. Product-focused HRIAs seek to provide guidance and identify considerations for product development.



01. End-to-End Encryption (“E2EE”)

E2EE keeps people and their personal communications safe from hackers, criminals and authoritarian regimes. In 2016, we implemented E2EE by default on WhatsApp and as an option on Messenger. In March 2019, we announced plans to extend this protection [by default across our messaging apps](#).

In October 2019, Meta commissioned Business for Social Responsibility (BSR) to undertake an HRIA of the expansion of E2EE. The subsequent two years of work took place during the reporting period. The [HRIA was published in full](#) in April 2022, accompanied by a [Meta response](#).

Salient Risks:

BSR noted a number of positive opportunities that E2EE would enable, including further protecting a diverse range of human rights such as: privacy, freedom of expression, protection against cybercrime threats, physical safety, freedom of belief and religious practices, and freedom from state-sponsored surveillance

and espionage. BSR also explained how human rights risks — such as child sexual abuse and exploitation, mis/disinformation, violent extremism, human trafficking — all operate independently of E2EE, and should be addressed separately, without undermining E2EE and focus on key marginalized communities.

Recommendations and Meta's Response:

The report [made a number of recommendations](#) to address risks and maximize the benefits of E2EE, divided among four categories of action: product, process, product policy and public policy; to wit:

- **Recommendations about specific products and features:** user reporting; UX and user testing; user education; languages; friction; opt-in account linking.
- **Recommendations for how Meta can detect and mitigate human rights risk:** process harm prevention strategies, such as metadata analysis; Machine Learning (ML) techniques for proactive detection; holistic child rights strategy; investigate client-side scanning techniques; assess impacts of cross-app communications.
- **Recommendations for product policy changes:** consistent privacy policies and improved transparency; define content standards; ML explainability; improved user reporting; user appeals transparency; grievance mechanisms.
- **Recommendations for how Meta should engage external stakeholders:** advocate for end-to-end encryption; engage policy makers; engage stakeholders; collaborate with researchers; quantify harms; collaborate with industry; train law enforcement.

Meta's [response](#) details our commitment to implementing 34 of the recommendations, partly implementing four, assessing the feasibility of another six and taking no further action on one — the pursuit of client-side scanning technologies.



The report concludes that client-side scanning technologies as they exist today would undermine the integrity of E2EE and disproportionately restrict people's privacy and other human rights.

Meta is committed to implementing the vast majority of the recommendations and is working diligently toward our plans for expanding E2EE as a means to help protect people and support their human rights.

C. Due Diligence Exercises

Meta uses rapid, decision-useful due diligence exercises to provide salient risk assessments for moments of crisis, immediate product development needs, and ongoing content policy decisions.

This work can build on the diligence exercise designating a country high priority, or be responsive to emerging needs. It can look back at moments of crisis to provide insights on improved processes going forward; or

run simultaneously to the crisis or product development to provide immediate, real-time feedback to inform decision-making. These assessments can be conducted in-house, or with a third-party. They are designed to run for shorter periods of time than our formal HRIAs (*i.e.*, weeks/months compared to years, on average). While their form can differ, their function is uniform: actionable, decision-useful, and immediate insights to guide the management of salient risks.



01. Ray-Ban Stories

Salient Risks:

In September 2021, in partnership with EssilorLuxottica, Meta introduced [Ray-Ban Stories](#), a first-generation pair of smart glasses that have cameras in the frames that can take photos and videos, which you can download to a companion app and then upload to social media.

Prior to launch, to mitigate human rights issues posed by people who may misuse such a device, Meta undertook an internal analysis

of the possible human rights implications of this product. Salient risks identified included: informed consent from bystanders, the safety of people wearing the device, as well as its effect on vulnerable groups that might be at increased risk of adverse effects from its operation, like women and children, human rights defenders, or minority groups. It also examined privacy risks and future, ongoing risks posed by any new technology, from visual search to the cloud storage of data.

Recommendations and Meta's Response:

The saliency assessment began by noting some positive human rights impact of hands-free and voice-controlled photo and video capabilities, including help for people with disabilities, or aiding accountability for human rights violators. It then noted if and where use of the device could implicate various human rights treaties, including things like: whether it could be seen by people with visual impairments, and if not whether that could violate the rights to dignity and privacy, or non-discrimination.

Its recommendations included other acts of human rights due diligence, including expanding user research to focus on vulnerable bystanders, like homeless communities, minority groups and human rights defenders, as well as recommendations to conduct further field-level focus groups to understand the product's impact *in situ*.

To mitigate these issues, the due diligence exercise recommended exploring several courses of action in consideration of bystanders, such as: developing an acceptable use policy, a do-not-disturb function, other signals for bystanders or the option to tag content as being from the device when it is shared.



D. Integrated Analytical Approaches

As part of our broader human rights due diligence efforts, we put into practice various integrated, analytical approaches that help us comprehend our most salient risks.



01. Countries At Risk

As we note in Section 02 of our [Human Rights Policy](#),

“We pay particular attention to the rights and needs of users from groups or populations that may be at heightened risk of becoming vulnerable or marginalized.”

In identified countries (and priority elections), we take a more customized comprehensive due diligence, risk mitigation and response approach — acting quickly to initiate the process to remove content that violates our policies and taking protective measures, including deploying country-specific support. Since 2018, this work has been led by teams with expertise in human rights, hate speech, women and child protection, misinformation, and polarization. Many have lived or worked in high-risk countries and speak relevant languages.

Beginning in 2019, we developed [a human rights due diligence process for reviewing and](#)

[prioritizing countries with the highest risk of offline harm and violence](#), every six months. To assess long-term conditions on the ground, a cross-functional team of data scientists, political scientists and regional experts review quantitative and qualitative data twice a year drawn from over 60 sources including [Varieties of Democracy \(V-Dem\)](#), [Uppsala Conflict Data Program](#), the [United States Holocaust Memorial Museum’s Early Warning Project](#), the [Armed Conflict Location & Event Data Project](#), and the World Bank. This data can include information regarding civic participation and human rights, societal tensions and violence, and the quality of relevant information ecosystems.

UNGP-Informed Factors to Determining Countries at Risk:

- **Long-term conditions and historical context:** We rely on regional experts, platform data and data from more than 60 sources to assess the long-term conditions on the ground.
- **How much the use of our products could potentially impact a context:** We prioritize countries based on a number of factors, including: where our apps have become most central to society, such as in countries where a larger share of people use our products; where there is potential for an increase in offline harms; and where social media adoption has grown.
- **Current events on the ground:** We give special consideration to discrete events that might magnify current societal problems, such as local risk or occurrence of atrocity crimes, polarizing elections, episodes of violence, and COVID-19 vaccination and transmission rates.

Guided by the above, we provide specialized support and focus teams on countries most at risk by working to drive human rights to the front of our product, policy, operations, and corporate decision-making. Our focus is threefold: remove content that violates our policies, respect people's right to free expression, and help to keep people safe both online and offline.

As part of our work, we [took the following actions](#):

- **Understand and engage with local contexts and communities** to ensure our enforcement accounts for local context. We expanded our global network of third-party fact-checkers, invested significant resources in more than 30 countries with active conflict or societal unrest, and together with UN partners and dozens of local and global NGOs, we developed programming to make online engagement safer.
- **Develop and revise policies to prohibit harmful content and behavior** by constantly evaluating and refining our policies to address evolving nuances of hate speech, identify groups at heightened risk of violence and [human rights abusers](#), or the potential for rumors and misinformation to contribute to physical harm, particularly in countries where ethnic and religious tensions are present.
- **Rapidly enhance our technology and enforcement to help keep our community safe** when the risk of impact is greater, *i.e.* in advance of elections and during periods of heightened unrest. For example, we may reduce or restrict distribution of content from accounts previously found to post content violating our policies, while our teams investigate. Also, we may offer additional features, *e.g.* providing people in Afghanistan following the Taliban takeover, the ability to lock their profile as an extra layer of privacy, security and protection.
- **Improve our moderation across languages** by adding more expertise.

In 2021, we hired content moderators in 12 new languages, including Haitian Creole, Kirundi, Tswana and Kinyarwanda.

i. Our Program In Action: Case Studies

Recent examples of our work in this area include our efforts in [Ethiopia](#), [Mexico](#), [Myanmar](#), [Afghanistan](#), and Haiti. We highlight two case studies which exemplify this work in action:

(1) Myanmar which held elections in 2020 and

experienced a military coup in early 2021; and (2) Ethiopia which, since November 2020, has experienced a brutal civil war characterized by severe human rights violations and massive humanitarian needs.

a. Myanmar

Evolving Challenge:

Ahead of the November 2020 elections in Myanmar, and given the Myanmar Military's (the "Tatmadaw") history of atrocities against the Rohingya people, Myanmar was identified as high priority and at high risk for human rights violations. Though Meta's efforts to protect platform integrity during the election were successful, the political climate substantially deteriorated, culminating in the 2021 coup d'état by the Tatmadaw that unleashed a wave of repression, attacks on peaceful demonstrations, and the silencing of dissenting voices.

Diligence:

A cross-functional, multi-disciplinary team came together to prepare for the elections; later, worked to identify the salient risks due to the coup, and took steps to help prevent and mitigate those risks.

Action:

Following the coup, the military government unleashed a wave of human rights violations against the people of Myanmar, protestors, human rights defenders, journalists, and others. We treated the situation as an emergency. Our Integrity Operations Center ran around the clock, bringing together subject matter experts from across the company, including Myanmar nationals with native language skills, to monitor and respond to threats in real time.

As a result, Meta took a series of [strong actions](#):

- Meta banned the Tatmadaw and military-controlled state and media entities from Facebook and Instagram, as well as ads from

military-linked commercial entities, and removed Pages, Groups and accounts representing military-controlled businesses. We based our actions on:

- Extensive documentation by the international community of these entities' direct role in funding the Tatmadaw's ongoing violence and human rights abuses, which also formed the basis for escalating sanctions imposed by the US, EU, and other governments.
- The UN Fact-Finding Mission on Myanmar's 2019 report on the economic interests of the Tatmadaw, in line with the UN Guiding Principles on Business and Human Rights.
- We reaffirmed Meta's ban on 20 military-linked individuals and organizations in 2018, including Commander-in-Chief Min Aung Hlaing, for their role in severe human rights violations; and the decision to remove at least six Coordinated Inauthentic Behavior networks run by the Tatmadaw from 2018 to 2020.
- Meta implemented a Myanmar-specific policy removing praise, support and advocacy of violence by Myanmar security forces against civilians and/or arrest of civilians by the military and security forces on our platform, under our [coordinating harm and publicizing crime policy](#).
- We offered a "Lock Your Profile" feature for Facebook, providing additional safety, privacy and security for our on-the-ground users.
- Since the coup, we disabled the Tatmadaw True News Information Team Page, and MRTV and MRTV Live Pages for continuing to violate our policies. We also reduced the distribution of content on at least 23 pages and profiles controlled and/or operated by the Tatmadaw so fewer people see them.

b. Ethiopia

Evolving Challenge:

As of this report's publication, Ethiopia is undergoing a conflict that began in November 2020 between the government of Ethiopia and forces in the northern Tigray region. As the turmoil spreads, the situation remains volatile. Ethiopia is an especially challenging environment to address these issues, in part because there are multiple languages spoken in the country. Meta and its platforms are key avenues for communication in Ethiopia.

Diligence:

Since early 2020, human rights due diligence for Ethiopia has included field-based information gathering and research; focused stakeholder engagement; and formal due diligence. In 2021, Meta commissioned forward-looking human rights due diligence to help identify existing and prospective human rights risks and inform internal decision-making. We engaged activists, human rights defenders, journalists, civil society groups throughout Ethiopia and the diaspora, adding to the insights gained regularly from human rights groups, NGOs, local civil society organizations, and other regional experts.

Action:

Using learnings from our due diligence, Meta increased efforts to tackle harmful content (e.g., violence and incitement, hate speech, coordinated inauthentic behavior, misinformation and harm), and addressed elections-related risks and content involving dangerous organizations and individuals. [We also:](#)

- **Improved and simplified our reporting tools** to make it easier for Ethiopian users to report possible violating content, so we can investigate.
- **Rolled out a [media literacy campaign](#)**, aimed at educating and informing people on how to detect potential false news, and ran billboard advertising campaigns across Addis Ababa, the first of its kind in Africa, focused on informing and educating people on how to stay safe online and use social media responsibly.



- **Partnered with the Center for African Leadership Studies** to implement “My Digital World,” a series of live webinars on topics such as online safety, privacy, digital citizenship, news and media literacy.
- Established **dedicated reporting channels** for specialized international and local human rights and civil society organizations to quickly review problematic content they identify for possible violations.
- **Launched a [Lock Your Profile](#) feature (as in Myanmar) for Facebook**, providing additional safety, privacy and security for our on-the-ground users.
- **Increased our investment in proactive detection technology** that helps identify violating content such as hate speech before it is reported, including adding in Amharic and Oromo capabilities.
- **Hired more Ethiopian nationals** with native language skills to monitor and respond to threats in real time and onboarded more Trusted Partners to provide contextualized advice.

02. Security Policies

Meta strives to combat coordinated inauthentic behavior, defined as coordinated efforts to manipulate public debate for a strategic goal where fake accounts are central to the operation. Such content on Meta's platforms could lead to salient human rights risks advanced through coordinated posts that:

- attack people on the basis of their opinion, beliefs, or protected characteristics;
- are connected to harm;
- depict violence;
- bully or harass users;
- encourage self-injury, misinformation or disinformation; or
- exacerbate conflict, corruption, and instability in conflict affected and high risk areas.

Meta performs ongoing diligence to review, assess, and ultimately help remove deceptive campaigns around the world. When we find campaigns that seek to mislead people about who they are and what they are doing using fake

accounts, we remove both the inauthentic and authentic accounts, Pages, and Groups directly involved. Using both automated and manual detection, Meta continues to remove accounts and Pages connected to networks that we took down in the past.

Under our UNGP and GNI commitments, this work is informed by stakeholders and performed with transparency. Since 2017, Meta has reported on over 150 influence operations and published information on threats we see from nation states, commercial firms or unattributed groups. We share information across the world with investigative journalists, government officials, independent researchers, and industry peers to better understand and expose internet-wide security risks, including ahead of critical elections. In 2021 alone, we removed 52 networks that originated in over 30 countries and expanded our beta research platform (containing over 100 data sets) to researchers studying influence operations.



i. Combating Emerging Threat Actors

The global threats we tackle have significantly evolved since we started sharing our findings about [Coordinated Inauthentic Behavior](#) in 2017. Adversarial networks are always looking for new ways to evade our policies. To account for this constantly shifting environment, we build our defenses to adapt and incorporate new layers of defense. Our goal over time is to make these behaviors more costly and difficult to hide, and less effective.

- **New Defenses against Coordinated Mass Harassment (“Brigading”), Abusive Reporting and Coordinated Violations of our Community Standards**

In December 2021, [we shared](#) our first report that brought together multiple network disruptions for distinct violations of our security policies: Coordinated Inauthentic Behavior and two new protocols — Brigading and Mass Reporting.



Brigading:

We remove any adversarial networks we find where people work together to mass comment, mass post or engage in other types of repetitive mass behaviors to harass others or silence them.

As detailed in the [December, 2021 Adversarial Threat Report](#), we removed a network of accounts that originated in Italy and France and targeted medical professionals, journalists, and elected officials with mass harassment. We also removed a network of accounts in Vietnam for violating our Inauthentic Behavior policy against mass

reporting by coordinating false reports against activists and others who publicly criticized the Vietnamese government in an attempt to have these users removed from Facebook.



Mass Reporting:

We remove any adversarial networks we find where people work together to mass-report an account or content to get it incorrectly taken down from our platform.

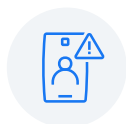
We are starting to remove coordinated violating networks when we find people — whether they use authentic or fake accounts — working together to violate or evade our [Community Standards](#).



ii. Protecting Users Against Espionage

Cyber espionage actors typically target people across the internet to collect intelligence, manipulate them into revealing information, and compromise their devices and accounts. When

we disrupt these operations, we take down their accounts, block their domains from being shared on our platform, and notify people who we believe were targeted by these malicious groups.



Our Campaign Against Surveillance for Hire

In December 2021, [we announced our action](#) against surveillance-for-hire entities operating on our platform, as part of a broader campaign effort that also included a call for further regulation and industry collaboration to counter this threat. Surveillance-for-hire entails commercial players duplicitously engaging in reconnaissance, engagement and exploitation of users on social media platforms, which can lead to salient human rights risks. During the reporting period, we took action against seven different surveillance-for-hire entities who indiscriminately targeted people in over 100 countries on behalf of their clients. These providers are based in China, Israel, India, and North Macedonia, [among others](#).



Broader Response to Abuse by Surveillance-For-Hire Groups

While cyber mercenaries claim that their services and surveillanceware are meant to focus only on criminals and terrorists, our own investigation, [independent researchers](#), our [industry peers](#) and [governments](#) have [demonstrated](#) that targeting includes journalists, dissidents, critics of authoritarian regimes, families of opposition and human rights activists. In fact, for platforms like ours, there is no scalable way to discern the purpose or legitimacy of such targeting. This is why we focus on enforcing against this behavior, regardless of who is behind it or who the target might be.

To support the work of law enforcement, we already have authorized [channels](#) where government agencies can submit lawful requests for information, rather than resorting to the surveillance-for-hire industry. These channels are designed to safeguard due process and we [report](#) the number and the origin of these requests publicly.

To [protect people](#) against cyber mercenaries operating across many platforms and national boundaries, we engage in a collective effort with other platforms, policymakers and civil society to counter the underlying market and its incentive structure.

Providing Access to Remedy

The third pillar of Meta's [Human Rights Policy](#) concerns the provision of remedies for human rights impacts.

Access to Remedy is a key component of the [UNGPs](#), with regard to both State and non-State-based grievance mechanisms. Affected rights-holders should be able to claim remedy without fear of victimization.

At Meta, we strive to provide pathways for stakeholders to report concerns, for Meta to review them, and for Meta to provide remedy and remediation consistent with UNGP 31. Meta maintains multiple grievance pathways, identified in the [Code of Conduct](#), in the

[Help Center](#) on platforms and apps, including an appeals process to the first-of-its-kind [Oversight Board](#). Anonymous reporting for Meta personnel is available through the [Whistleblower and Complaint Policy](#).

We seek to provide meaningful transparency, with routine publication of policies with human rights implications, information related to how content determinations can be appealed, and decisions of, and responses to, the Oversight Board. When content is removed, we notify users and clearly identify pathways for content removal decisions to be appealed. Data regarding removal decisions and other aspects of policy enforcement are published through the [Transparency Center](#).



A. The Oversight Board

After extensive global consultation, Meta founded the [Oversight Board](#) in 2020. The Oversight Board was born, in part, to respond to concerns that Meta needed an independent third party to adjudicate violations of its rules, provide policy recommendations and address access to remedy.

The Board [consists](#) of experts from a variety of cultural and professional backgrounds, chosen because they are “experienced at deliberating thoughtfully and collegially, skilled at making and explaining decisions based on a set of policies or principles, and familiar with digital content and governance.”¹⁴

As its [charter makes clear](#), the purpose of the board is to: “[p]rotect free expression by making principled, independent decisions about important pieces of content and by issuing policy advisory opinions on [Meta’s] content policies.”

The independence of the Oversight Board is strongly protected by its Charter, its Bylaws, its members, and its funding.

Prior to its launch, Business for Social Responsibility completed a [Human Rights Review](#) of the Oversight Board. This review made a series of recommendations, including prioritizing cases that present the most severe human rights harms, using the UNGP’s scope, scale and remediability criteria. The review also recommended that the human rights impacts of each case be reviewed, and that the human rights principles of legitimacy, necessity, and proportionality be used. One year later, a follow up [human rights review](#) by Business for Social Responsibility found that “the Oversight Board’s governing documents provide an increasingly valuable framework for [taking a] human rights-based approach to content decisions.”

The Oversight Board plays a critical, industry-leading role in the fulfillment of Meta’s responsibility under the UNGPs to provide “Access to Remedy,” which requires business

enterprises to “establish or participate in effective operational-level grievance mechanisms for individuals and communities who may be adversely impacted [by their operations].” This not only makes it possible to “identify adverse human rights impacts as a part of an enterprise’s ongoing human rights due diligence,” it also “makes it possible for grievances, once identified, to be addressed and for adverse impacts to be remediated early and directly by the business enterprise, thereby preventing harms from compounding and grievances from escalating.”

For its part, the Meta Human Rights Policy Team supported this project by working with colleagues on initial due diligence, the Charter, member selection, initial referrals, board education, training of Meta staff for interacting with the board, due diligence on transparency and follow up to board recommendations.

Following Oversight Board decisions, Meta started the process of reinstating [identical content with parallel context](#) in the following cases: [Uyghur Muslims](#), [Hydroxychloroquine](#), [Azithromycin](#) and [COVID-19](#), and a [Nazi Quote](#). These actions affect not only content previously posted on Facebook and Instagram but also future content. For cases where the board upholds our judgment, we continue to ensure identical content with parallel context remains either up or down, in line with the board’s decision.

The Oversight Board has [directly called for transparency](#) when circumstances require it, and [brought in for discussion](#) people critical of the company.

14. <https://oversightboard.com/meet-the-board/>

In another example of its focus on transparency, when Meta referred to the Board a question regarding its restriction of then-US President Donald Trump’s access to posting content on his Facebook page and Instagram account, the Board observed that Meta failed to mention a “cross-check” system, and that, “Given that the referral included a specific policy question about account-level enforcement for political leaders, many of whom the Board believes were covered by cross-check, this omission is not acceptable.”

The Board, via its case decisions and policy recommendations, is defining what a human rights-based approach to content means in practice, to a level of detail, practical application and intellectual rigor, that did not previously exist.

The Oversight Board publishes a quarterly [transparency report](#) highlighting its extensive efforts adjudicating the most difficult and pressing speech issues facing Meta. Cases before the board include those submitted by Meta, users or others.



B. Code of Conduct

[Meta’s Code of Conduct](#) sets out guidelines for business conduct required of all Meta personnel, and is accompanied by multiple channels to report concerns or violations, and potential remedies. It helps embed human rights into our overall governance by including respect for human rights

as a core part of our expectations for all Meta personnel.

The ability to make anonymous reports is covered by Meta’s [Whistleblower and Complaint Policy](#). Retaliation or intimidation against individuals reporting issues in good faith is strictly forbidden.

Protecting Human Rights Defenders



The fourth pillar of our [Human Rights Policy](#) is dedicated to protecting Human Rights Defenders, as defined by the [UN Declaration on Human Rights Defenders](#), recognizing them as a priority group for support and protection.

We commissioned internal due diligence in 2019–20, the goal of which was to take a first step toward designing a human rights defender program, mapping existing work, identifying gaps and opportunities, and suggesting possible next steps. We then established an internal Human Rights Defender working group, which has significantly improved communication and information-sharing across teams and helped to identify gaps, priorities, and potential interventions to better support human rights defenders. These interventions include steps related to safety, security, research, policy development, and product or operational support.

We're building on [our existing work to protect defenders' accounts](#) — efforts that include [combatting advanced threat actors](#) targeting them, protecting them from incorrect content removals using [Cross-Check](#), offering advanced [security options](#), taking steps to thwart unauthorized access to the accounts of defenders who are arrested or detained, and partnering with human rights organizations on outreach and training.

One important workstream has been our work on civic actors, who may face higher risks due to their contributions to civic affairs, either online or offline. These risks include threats to privacy, reputation, well-being or physical safety. Interventions applied for civic actors include harassment monitoring, enhanced account security and impersonation takedowns.

WhatsApp's end-to-end encryption is [widely recognized](#) to be essential for the protection of human rights defenders and journalists — as well for racial and religious minorities, LGBTQ+ communities and other disadvantaged groups.

Engagement with human rights defenders (and related UN actors) has increased significantly. We are driving regular engagement at the local and global levels to inform our understanding of related risks, mitigations, and challenges. Our engagements have included [digital security training for activists](#) in a number of countries.



A good example of innovative work in the Africa region has been the creation of the Sub-Saharan Africa (SSA) Women's Working Group, which was convened in 2021 for women activists across SSA working on issues that intersect with gender equality and digital spaces. Participants represented grassroots, regional and international organisations covering issues ranging from online gender-based-violence, feminist rapid-response services, online safety for the LGBTQ+ community, tech-feminism and tech law & governance.

The working group helped us improve policies and products affecting women users on Facebook platforms and include these voices in ongoing work that requires engagement for content policies and other platform developments and products affecting this constituency.

When Meta launched our [Human Rights Policy](#) in March of 2021, we also committed to [create a fund](#) that will give offline assistance to human rights defenders facing critical threats and support new digital security efforts.



This fund was formally launched in [November 2021](#), as the Human Rights Defender Fund and Journalist Safety Initiative, in partnership with [Civil Rights Defenders](#) and [International Center for Journalists](#) and is being piloted in the Asia-Pacific region.

The fund supports devices and security technologies, temporary relocation, and emergency legal and safety assistance, as well as mental health support to human rights defenders and journalists. It also enables the development of peer-to-peer support programs and relevant mental health practitioner networks. The fund is being distributed over a two-year period, and managed independently.

We have made good progress in integrating the protection of human rights defenders into our responses in countries at-risk and will continue to look to develop and roll out new policies and products in more regions — especially those identified as most critical for the well-being of human rights defenders and other civic actors.

Governance, Oversight, Accountability

The fifth core pillar of our Human Rights Policy reflects our commitment to UNGP 16 to embed respect for human rights throughout our enterprise.



A. Governance Model

Meta's [Human Rights Policy](#) and its commitments to human rights are enterprise-wide. Nick Clegg (President, Global Affairs) and Jennifer Newstead (Chief Legal Officer) oversee policy implementation. Salient risks are reported periodically to the Audit and Risk Oversight Committee of the Meta Board of Directors.

The global Human Rights Policy Team drives application of the Policy, setting the company's human rights agenda, identifying priorities, shaping strategy, inspiring and informing colleagues, and engaging with human rights activists and experts. This team was created in 2019, with the appointment of Miranda Sissons as its inaugural director and by the end of 2021 had four full-time staff, with several more positions onboarding in 2022.¹⁵

In a company of tens of thousands of employees and contingent workers, and with billions of people using our apps and services, a key challenge for the Human Rights Policy program is to ensure that we leverage our expertise and resources for maximum impact.

We seek to ensure that salient human rights issues are directed to the most relevant decision-making forum. The Human Rights Policy Team functions in an umbrella governance structure: consulting with and convening the functional units implementing human rights-informed policy and product changes.

15. As of March 2022 the global Meta Human Rights Policy Team consists of 9 full time employees.

Teams working on community operations, law enforcement and security, responsible innovation, product integrity and public policy (among others) regularly identify, prevent, or mitigate salient human rights risks. Similarly, teams focused on civil rights, Africa, the Middle East and Turkey, sustainability, International Organizations, and Stakeholder Engagement also work to implement policy goals.

Key activities to ensure proper governance over human rights responsibilities include:

- Convening and driving the work of various cross-functional, inter-departmental working groups dedicated to the implementation of the Human Rights Policy;
- Joining cross-functional, inter-departmental working groups to provide human rights perspective and principles to their work;

- Driving human rights due diligence processes with key product and public policy colleagues;
- Participating actively in the at-risk country work, ensuring that human rights considerations are effectively prioritized;
- Coordinating work on human rights defender protection;
- Representing Meta at various human rights fora including the GNI, [B-Tech](#), and with key UN and NGO human rights agencies.

[Meta's Board of Directors' Audit & Risk Oversight Committee's charter](#) provides for the committee's oversight of programs, policies and risks related to social responsibility, which includes oversight of human rights-related risks. Beginning in 2021, the Director of the Human Rights Policy Team updates the Audit and Risk Oversight Committee at least annually.

In the course of the last two years, Meta's teams have worked across a wide range of human rights issues, including but not limited to:

- Business and Human Rights;
- Freedom of Expression (particularly proportionality);
- Privacy;
- Freedom of Assembly;
- Political participation;
- Civic engagement;
- Non-discrimination;
- International humanitarian law;
- International criminal law;
- Transitional justice processes, particularly process design for truth seeking and reparations initiatives;
- Agile software development processes;
- Atrocity crime prevention;
- Child rights;
- Regional human rights practices and instruments;
- Modern slavery;
- Lawful use of lethal force and law enforcement principles;
- Right to health;
- Machine learning techniques;
- The San Remo Principles;
- Open source evidence investigation;
- Quantitative and qualitative research processes and indicators;
- Comparative law;
- Know your customer (KYC) processes; and
- Adult learning practices.

B. Stakeholder Engagement



Our [Human Rights Policy](#) is founded on deep engagement with rights-holders and other stakeholders.¹⁶ Stakeholder Engagement is thus relevant to every pillar of the policy.

Stakeholder engagement is a company-wide effort, integrated into several company processes. When developing the Human Rights Policy, for example, the Human Rights Policy Team engaged extensively with human rights experts and activists, including activist investors. Team members are in regular contact with human rights organizations, activists, academics and others to hear concerns and criticisms, seek advice and brief on ongoing developments.

As described in earlier sections of this report, such engagement includes consultations and co-design with human rights defenders on measures to protect account security, engagement on coordinated inauthentic behavior and espionage, roundtables with regional civil society actors ahead of elections, and much more.

Stakeholder engagement is also an important part of Meta's human rights due diligence process. It enables us to identify and understand the human rights impact of our products and services as

well as develop appropriate mitigation strategies, including fair and transparent policies. During the reporting period, we also sought to pioneer new methodologies of direct rights holder consultation. For example, we combined a large scale quantitative survey of rights holders with traditional methods of stakeholder consultation in our human rights due diligence in the [Philippines](#).

In another important example, the Content Policy Stakeholder Engagement team strengthens our content policies by bringing global knowledge and feedback into the policy development process. This team builds relationships with NGOs, academics and other thought leaders, guided by our [core principles](#) of inclusivity, expertise, and transparency.

In all, the team covers the following policy areas: Community Standards, advertising policies, product policies, misinformation, content ranking, and news integrity.

[Work with stakeholders crafting our policies](#) includes:

16. By "stakeholders" we mean rights-holders, all organizations, and individuals who are impacted by, and therefore have a stake in, Meta's policies

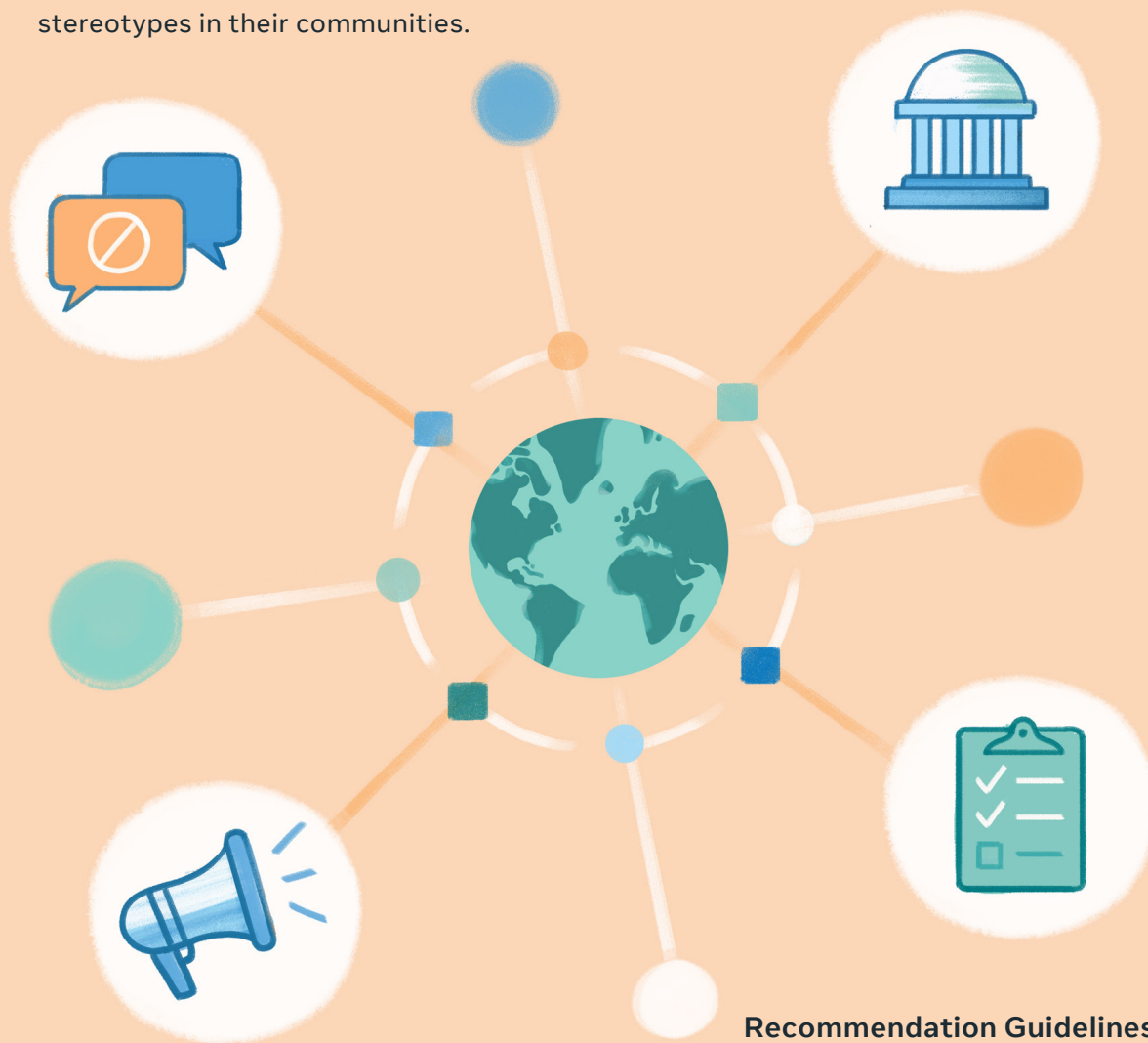
Engagement on Stereotypes

We created a policy on harmful stereotypes. As part of our approach to [hate speech and harmful stereotypes](#), we convened several academic roundtables and connected with external stakeholders who helped us identify the risks associated with certain stereotypes in their communities.

Understanding State Media

To inform our policy definition and criteria for [state media](#), we consulted with more than 65 global experts specializing in media, governance and human rights.

This input was crucial to helping us understand the different ways and degrees to which governments exert editorial control over certain media entities.



Guidance on “Public Figures”

We adopted a definition of “public figure” for our [bullying and harassment](#) policies based on engagement with academics, civil society organizations, free expression experts, human rights defenders, journalists and creators/influencers.

Recommendation Guidelines

In developing our Recommendations Guidelines for the [Facebook app](#) and [Instagram](#), we consulted with 50 leading experts who specialize in recommender systems, expression, safety and digital rights.

C. Program Spotlight: Trusted Partners

Meta's Trusted Partner (TP) program is an escalation channel made up of over 400 key non-governmental, not-for-profit, humanitarian, and international organizations in 113 countries who report content, accounts, and behavior that we review in context.

Trusted Partners are critical allies in our efforts to understand the local impact of our services, stay abreast of emerging trends, correct mistakes made in scaled review, and improve policies and enforcement in specific policy areas such as misinformation and harm. Their reports provide valuable insights, data points, and edge cases, which help inform policy development processes and help keep our users safe both on and offline.

In the past, TPs have alerted us to developments and risks that our internal monitoring efforts had missed. In Myanmar, for example, TPs provided us with additional context about the recent coup and online risks that emerged in its aftermath. In Ethiopia, our TPs helped us combat harmful misinformation and hate speech stoking inter-ethnic violence. In Afghanistan, our civil society partners helped us to mitigate the targeting of international aid workers and other vulnerable persons in our platforms.

The Trusted Partner program also provides a vehicle for amplifying the voices of marginalized communities, incorporating relevant information about the risks they face on our platforms, and providing support to under-represented and disproportionately-impacted groups.



D. Engagement with International Organizations

Part of our accountability process under our governance pillar is transparent, collaborative, and reciprocal engagement with the UN. Our International Institutions and Relations team supports engagement across the UN and intergovernmental organization system, including dialogue with UN human rights actors, which continues to broaden and deepen both at global and country level.

As long-running human rights issues continue to manifest online, we see this collaborative discourse with the UN human rights system as particularly valuable — this includes thematically focused actors such as UN Human Rights Council-mandated Special Rapporteurs, but also broader dialogue on tech and human rights across the UN Secretariat and with specialized agencies such as UNESCO.

- The International Institutions and Relations and Human Rights Policy Teams have met regularly with the UN Office of the High Commissioner for Human Rights since 2018, and we have been active members of the [B-Tech Project](#).
- In addition, we meet with various entities of the UN system including the Office of the Special Advisor for the Prevention of Genocide, as well as UN Special Rapporteurs on: Freedom of Expression, Human Rights Defenders, Protection of Minorities, Myanmar, Racism and Racial Discrimination and Freedom of Religion and Belief to discuss the intersection of human rights and technology, and its relevance to the actions of the UN system across peace and security, human rights, and humanitarian agendas.
- We also work with UN entities as trusted partners at both global and country level — including Haiti, Bosnia, and Ethiopia. We work with them to mitigate content risks and to better understand the links between on-platform conduct and off-line harms. We are collaborating on issues of racial justice, the implementation of the UNGPs and the critical importance of end-to-end encryption. We are contributing our expertise and experience to a variety of human rights reporting and norm

development processes; and have briefed and shared insights from our at risk countries with multiple relevant groups.

- Meta briefed the UN Security Council's October 28, 2021 [Arria Formula Meeting](#) on hate speech and social media.
- A key part of our [engagement](#) with the UN system and other multilateral actors has been to encourage a strong defense of an open and unified internet, with an emphasis on the many negative human rights impacts of internet fragmentation. Meta is also part of the [Internet Fragmentation](#) group.
- Meta has been represented in several human rights and hate speech related panels, including by senior executives, to further discourse and thinking around tackling hate speech and misinformation.
- As part of UNESCO's [Social Media 4 Peace](#) project, we have worked with UN representatives in Bosnia and Kenya to discuss with civil society and government stakeholders, ideas of how social media platforms can promote peace.
- Since June 2021, we have voluntarily disclosed 16 productions with organic content from hundreds of Pages, and Basic Subscriber information for hundreds of Pages and Profiles, to the Independent Investigative Mechanism for Myanmar (IIMM) which is seeking accountability for the most serious international crimes and violations of international law committed in Myanmar.
- In September 2021, Meta joined the [UN Global Compact](#), the world's largest corporate sustainability initiative. This marked an important milestone in Meta's commitment to human rights, labor, environment and anti-corruption. To take actions that advance societal goals, Meta has committed to the Global Compact's [10 Principles](#) on Human Rights, Labor, Environment, and Anti-Corruption.

A Final Note

“The saddest aspect of life right now is that science gathers knowledge faster than society gathers wisdom.”

ISAAC ASIMOV



Isaac Asimov once wrote, “The saddest aspect of life right now is that science gathers knowledge faster than society gathers wisdom.”

While this report was being written, a war in Ukraine was fiercely waged, precipitated by a Russian invasion. As with the Arab Spring of 2011, content on social media is seen to have a profound role in shaping how this conflict is perceived, and potentially, its dynamics.

This conflict, along with other crises in countries across the world, underscores how important it is for Meta and other companies to live up to their human rights commitments — for wisdom to always accompany knowledge.

We hope this report gives detailed insight into the work Meta has done to identify, mitigate, and prevent human rights risks from 2020 — 2021. We look forward to more work, and more reporting, to come.