

UNIVERSIDADE DO EXTREMO SUL CATARINENSE - UNESC

PROGRAMA DE PÓS-GRADUAÇÃO EM DIREITO

MESTRADO EM DIREITO

FERNANDA DA LUZ BERNA

**OS NOVOS DIREITOS NA SOCIEDADE DA INFORMAÇÃO E OS LIMITES À
LIBERDADE DE EXPRESSÃO NAS MÍDIAS SOCIAIS: UMA ANÁLISE DO CASO
DONALD TRUMP NO *OVERSIGHT BOARD***

CRICIÚMA

2023

FERNANDA DA LUZ BERNA

**OS NOVOS DIREITOS NA SOCIEDADE DA INFORMAÇÃO E OS LIMITES À
LIBERDADE DE EXPRESSÃO NAS MÍDIAS SOCIAIS: UMA ANÁLISE DO CASO
DONALD TRUMP NO *OVERSIGHT BOARD***

Dissertação apresentada ao Programa de Pós-Graduação em Direito – Mestrado – Área de Concentração em Direitos Humanos e Sociedade, Linha de Pesquisa em Direitos Humanos, Cidadania e Novos Direitos da Universidade do Extremo Sul Catarinense - UNESC, como requisito parcial para a obtenção do título de Mestre em Direito.

Orientador: Prof. Dr. Gustavo Silveira Borges.

CRICIÚMA

2023

Dados Internacionais de Catalogação na Publicação

B104n Berna, Fernanda da Luz.

Os novos direitos na sociedade da informação e os limites à liberdade de expressão nas mídias sociais : uma análise do caso Donald Trump no *Oversight Board* / Fernanda da Luz Berna. - 2023.
105 p. ; il.

Dissertação (Mestrado) - Universidade do Extremo Sul Catarinense, Programa de Pós-Graduação em Direito, Criciúma, 2023.

Orientação: Gustavo Silveira Borges.

1. Liberdade de expressão. 2. Mídia social. 3. Direitos humanos 4. Comitê de supervisão. 5. Sociedade da informação. I. Título.

CDD. 22. ed.
341.2727

Bibliotecária Eliziane de Lucca Alosilla - CRB 14/1101
Biblioteca Central Prof. Eurico Back - UNESC

FERNANDA DA LUZ BERNA

**“SOCIEDADE DA INFORMAÇÃO, LIBERDADE DE EXPRESSÃO E LIMITES DOS
PADRÕES DA COMUNIDADE NAS MÍDIAS SOCIAIS: ANÁLISE DO CASO DONALD
TRUMP NO OVERSIGHT BOARD”**

Esta dissertação foi julgada e aprovada para obtenção do Grau de Mestre em Direito no Programa de Pós-Graduação em Direito da Universidade do Extremo Sul Catarinense.

Criciúma, 29 de março de 2023.

BANCA EXAMINADORA


Prof. Dr. Gustavo Silveira Borges
(Presidente e Orientador (a) – UNESC)

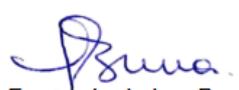

Prof. Dr. Daniel Ribeiro Preve
(Membro – PPGD/UNESC)

Yduan de Oliveira May
(Membro Suplente – PPGD/UNESC)


Prof. Dr. Reginaldo de Souza Vieira
Coordenador PPGD

Documento assinado digitalmente
 FILIPE JOSE MEDONAFFONSO
Data: 31/08/2023 14:44:21-0300
Verifique em <https://validar.itl.gov.br>

Prof. Dr. Filipe José Medon Affonso
(Membro Externo – FGV/RJ)


Fernanda da Luz Berna
(Mestrando(a))

Dedico esse estudo a todos aqueles que, de alguma forma, me ajudaram a chegar até aqui e que não soltaram a minha mão, mesmo quando eu quis soltar.

AGRADECIMENTOS

Em pensar que esta parte da construção da minha dissertação fosse a mais fácil. Muito foi pensado e repensado para escrever essas linhas de agradecimentos, pois muitos foram os desafios e obstáculos enfrentados e quase todos superados no transcorrer dos últimos anos.

Ter tido a oportunidade de ingressar no mestrado foi e está sendo a realização de um sonho que há muito tempo habilita no meu íntimo e que por muito tempo estava adormecido. Para quem não me conhece, ao invés de fazer o ensino médio, optei por cursar o magistério – em alguns locais chamado de Curso Normal – foi desde então que descobri minha vocação para as salas de aula. Desde então tenho comigo preceitos de Paulo Freire que acreditava que a educação não transforma o mundo, mas sim as pessoas e são essas pessoas que transformam o mundo.

Acredito que tudo tem um propósito, por isso, agradeço primeiramente a Deus por ter direcionado o meu caminho até aqui e me fortalecido durante essa jornada. Sem a fé de que tudo vai passar, que tudo tem o momento certo para acontecer e, principalmente, a fé de que o Senhor não me abandonou e me ajudou a superar todas as dificuldades encontradas nesse percurso.

Agradeço a minha família, meu marido, meus pais e familiares que estiveram presentes nessa etapa da minha vida, sempre vibrando e torcendo por mim, me dando apoio nos momentos em que eu achei que não iria conseguir, por sempre me darem todo o amor e carinho nos meus momentos de fraqueza e incerteza. Com toda a minha convicção, vocês foram um dos meus alicerces para me manter em pé e dando continuidade aos meus estudos.

Não posso deixar de agradecer ao meu orientador, professor Dr. Gustavo Borges, que sempre foi muito atencioso comigo, me auxiliando sempre que eu precisava e me incentivando durante essa jornada. Agradeço aos ensinamentos, aos conselhos e, inclusive, às críticas, pois nem sempre conseguimos acertar, mas é sempre bom contar com quem quer nos ver vencer e aperfeiçoar. Obrigada por todo apoio, professor, principalmente nos momentos que mais precisei de compreensão, você sempre foi muito acolhedor.

Preciso, ainda, agradecer meus colegas do mestrado, que passaram comigo essa jornada árdua, especialmente à Patrícia Longaretti, à Isabela Coppetti e ao Leonardo Schilling, que vou levar para sempre no meu coração. Obrigada por todo

o suporte e auxílio que vocês me deram em todos os momentos, vocês foram muito importantes para mim e para a construção dessa dissertação. Fico muito feliz e grata de ter encontrado vocês nesse caminho e, também, em saber que sempre pude contar com vocês para tudo.

Agradeço à Unesc e ao Programa de Pós-Graduação em Direito pela oportunidade que me foi dada, pois, sem sombra de dúvidas, foi uma experiência única na minha vida, a qual mudou muito o meu olhar para mundo, a minha perspectiva sobre diversos assuntos, os quais nunca tinha escutado falar antes de ingressar no mestrado. De fato, ter feito parte do PPGD da Unesc me mudou e sei que irá mudar muito a minha vida.

Por fim, agradeço à Secretaria do Estado de Santa Catarina que através da bolsa de mestrado, pela Uniedu, me proporcionou cursar o mestrado com mais tranquilidade em relação às questões financeiras. Sem esse recurso que me foi dado, não seria possível eu ter realizado esse sonho.

*Apesar de toda a conversa a respeito de apatia
e desmotivação da sociedade — parte dela
justificada, parte não —, permaneço otimista.*

Matthew D'Ancona

RESUMO

A dissertação tem por objetivo geral estudar de que maneira o *Oversight Board* pode garantir a proteção do direito humano à liberdade de expressão nas mídias sociais. A partir do objetivo geral, elaboraram-se os seguintes objetivos específicos: a) descrever as revoluções da sociedade da informação até o surgimento das mídias sociais e a necessidade de reconhecimento dos novos direitos da Era Digital; b) compreender aplicação do direito à liberdade de expressão e identificar seus limites legais por meio dos padrões internacionais de direitos humanos, padrões da comunidade do *Facebook* e a legislação norte-americana; e, por fim, c) analisar o julgamento da suspensão de contas do ex-presidente dos Estados Unidos, Donald Trump, através das premissas e conceitos estudados sobre liberdade de expressão, seus limites e o funcionamento do *Oversight Board*. O problema de pesquisa consiste em: de que maneira o *Oversight Board* pode garantir a proteção do direito humano à liberdade de expressão, a partir do estudo dos Padrões da Comunidade do *Facebook*, os *standarts* Internacionais de direitos humanos sobre liberdade de expressão e o direito norte-americano, tendo em conta a análise do caso referente à suspensão das contas do *Facebook* e *Instagram* do ex-presidente dos Estados Unidos, Donald Trump? Por sua vez, a pesquisa adotou o método de abordagem dedutivo; como método de procedimento, o monográfico e a técnica de pesquisa bibliográfica. Dentre os resultados e conclusões, destacaram-se: embora o *Board* não possua uma grande segurança jurídica, estando à mercê de seu criador para acatar ou não suas deliberações e sugestões, ressalta-se que a eleição de diretrizes do direito internacional voltados aos direitos humanos, dá ao “órgão julgador privado” uma aparência de maior neutralidade em suas decisões. A aplicação dos princípios internacionais de legalidade, legitimidade, necessidade e proporcionalidade na limitação à liberdade de expressão, auxiliam os julgadores a verificarem no caso concreto o tipo de medida que mais se adequa na moderação de conteúdo, o que seria aplicado em qualquer Corte internacional de direitos humanos, caso fossem designadas a este papel. Partindo desses princípios em relação ao caso analisado, o Comitê teve que mensurar não somente a violação que seu conteúdo havia incorrido, mas as proporções que seu discurso tomou para além dos limites do ciberespaço, os danos presente e futuros (médio e longo prazo) que ocorreriam se o conteúdo e seu interlocutor se mantivessem ativos na rede social. O que não se trata de medida de censura, pois muito embora a suspensão por tempo indeterminado foi evidentemente desproporcional, a medida em si foi necessária na contenção de prejuízos à segurança da população e, também, da democracia norte-americana. O Comitê pode não ser a solução do futuro para esse grande problema da Era Digital, mas, sem sombra de dúvidas, é um caminho que será seguido, remodelado e replicado para melhor resguardar o direito fundamental à liberdade de expressão em detrimento à moderação de conteúdo nas mídias sociais,

Palavras-chave: comitê de supervisão; liberdade de expressão; mídias sociais; moderação de conteúdo.

ABSTRACT

The dissertation aims as a general objective to study how the Oversight Board can guarantee the protection of the human right to freedom of expression on social media. Based on the general objective, the following specific objectives were developed: a) describe the revolutions in the information society until the emergence of social media and the need to recognize the new rights of the Digital Era; b) understand the application of the right to freedom of expression and identify its legal limits through international human rights standards, Facebook community standards and North American legislation; and, finally, c) analyze the trial regarding the suspension of accounts of the former president of the United States, Donald Trump, through the premises and concepts studied about freedom of expression, its limits and the functioning of the Oversight Board. The research problem consists of: how can the Oversight Board guarantee the protection of the human right to freedom of expression, based on the study of the Facebook Community Standards, the International human rights standards on freedom of expression and the right American, taking into account the analysis of the case regarding the suspension of the Facebook and Instagram accounts of the former president of the United States, Donald Trump? In turn, the research adopted the deductive approach method; as a method of procedure, the monograph and the bibliographic research technique. Among the results and conclusions, the following stood out: although the Board does not have great legal security, being at the mercy of its creator to accept or not its deliberations and suggestions, it is noteworthy that the election of international law guidelines focused on human rights, gives the “private court” an appearance of greater neutrality in its decisions. The application of international principles of legality, legitimacy, necessity and proportionality in limiting freedom of expression, helps judges to verify in the specific case the type of measure that is most appropriate in moderating content, which would be applied in any international human rights court, if they were assigned this role. Based on these principles in relation to the case analyzed, the Board had to measure not only the violation that its content had incurred, but the proportions that its speech took beyond the limits of cyberspace, the present and future damages (medium and long term) that would occur if the content and its interlocutor remained active on the social network. This is not a censorship measure, because although the indefinite suspension was evidently disproportionate, the measure itself was necessary to contain damage to the safety of the population and also to North American democracy. The Oversight Board may not be the future solution to this major problem of the Digital Age, but, there is no doubt, it is a path that will be followed, remodeled and replicated to better protect the fundamental right to freedom of expression to the detriment of content moderation on social media.

Keywords: content moderation, freedom of expression, oversight board, social media.

LISTA DE ABREVIATURAS E SIGLAS

| | |
|--------|---|
| BDTD | Biblioteca Digital Brasileira de Teses e Dissertações |
| CADH | Convenção Americana de Direitos Humanos |
| CDA | <i>Communications Decency Act</i> |
| CIDH | Comissão Interamericana de Direitos Humanos |
| DSA | <i>Digital Services Act</i> |
| DUDH | Declaração Universal dos Direitos Humanos |
| EFF | <i>Electronic Frontier Foundation</i> |
| EUA | Estados Unidos da América |
| IA | Inteligência artificial |
| KKK | <i>Ku Klux Klan</i> |
| MIT | <i>Massachusetts Institute of Technology</i> |
| NetzDG | <i>Netzwerkdurchsetzungsgesetz</i> |
| OEA | Organização dos Estados Americanos |
| ONU | Organização das Nações Unidas |
| PIDCP | Pacto Internacional dos Direitos Civis e Políticos |
| SciELO | <i>Scientific Electronic Library Online</i> |
| TIC | Tecnologias da Informação e Comunicação |
| UNESC | Universidade do Extremo-Sul Catarinense |

LISTA DE FIGURAS

| | |
|--|----|
| Figura 1 – Estimativa de submissão de casos ao Comitê por ano | 77 |
| Figura 2 – <i>Tweet</i> de Trump “ <i>Stop the count</i> ” | 81 |
| Figura 3 – <i>Tweet</i> de Trump afirmando fraude nas eleições | 81 |
| Figura 4 – <i>Tweet</i> de Trump afirmando que havia vencido as eleições | 82 |
| Figura 5 – <i>Tweet</i> de Trump afirmando que houve trapaça nas eleições | 82 |
| Figura 6 – <i>Tweet</i> de Trump informando que a “batalha” estava começando | 83 |
| Figura 7 – <i>Tweet</i> de Trump afirmando corrupção nas eleições | 83 |
| Figura 8 - <i>Tweet</i> de Trump chamando seus eleitores para a capital no dia 06/01 | 84 |
| Figura 9 – <i>Tweet</i> de Trump convocando seus eleitores para o protesto em 06/01 | 84 |

SUMÁRIO

| | |
|--|------------|
| 1 INTRODUÇÃO | 13 |
| 2 A SOCIEDADE CONTEMPORÂNEA, MÍDIAS SOCIAIS E A NOVA ÉGIDE DOS DIREITOS HUMANOS NA SOCIEDADE DA INFORMAÇÃO..... | 19 |
| 2.1 A sociedade da informação e a vida <i>onlife</i> | 19 |
| 2.2 Do surgimento das mídias sociais à utilização das plataformas para disseminar atos de violação de direitos humanos..... | 27 |
| 2.3 Da necessidade de novas configurações para a proteção dos direitos humanos na Era Digital | 36 |
| 3 A LIBERDADE DE EXPRESSÃO, SEUS LIMITES E OS PADRÕES DE COMUNIDADE DO META | 47 |
| 3.1 Os padrões internacionais sobre liberdade de expressão e garantia dos direitos humanos na moderação de conteúdo nas mídias sociais | 49 |
| 3.2 A autorregulação dos provedores intermediários, Termos de Uso e os padrões da comunidade do <i>Meta Inc.</i> | 58 |
| 3.3 Os limites da liberdade de expressão nos Estados Unidos | 65 |
| 4 ANÁLISE DO JULGAMENTO DAS CONTAS DE DONALD TRUMP PELO OVERSIGHT BOARD..... | 77 |
| 4.1 O <i>Oversight Board</i> : criação, composição e funcionamento | 77 |
| 4.2 Relato sobre os atos que antecederam a suspensão das contas de Donald Trump | 82 |
| 4.3 A análise do julgamento do caso 2021-001-FB-FBR | 91 |
| CONCLUSÃO | 96 |
| REFERÊNCIAS..... | 101 |

1 INTRODUÇÃO

Com o passar das revoluções da sociedade e a evolução das tecnologias de informação e comunicação, as redes sociais surgiram como um meio facilitador de contato social, possibilitando a interação de pessoas ao redor de todo o mundo e a aproximação de grupos sociais de acordo com suas preferências e escolhas em comum.

Nesse sentido, a razão pela qual um estudo sobre a temática proposta é de grande relevância, se dá pelo fato que o mundo está cada vez mais conectado não somente à internet, mas também às mídias sociais. Logo, a cada crescimento dessa população digital e sua fácil aderência à utilização dessas plataformas, como as do *Facebook* e *Instagram*, faz com que as postagens envolvendo conteúdos violadores de direitos humanos circulem de uma maneira muito mais rápida e tenham uma repercussão maior.

No início da presente pesquisa, foi levantado que, de acordo com o relatório *Digital 2022*, publicado pela *We are social* em parceria com a *Hootsuite*, o alcance da conexão à Internet mundial já atingia 62,5% dos habitantes do planeta, ou seja, das 7,91 bilhões de pessoas existentes, 4,95 bilhões de indivíduos possuem acesso à Internet, enquanto 4,62 bilhões são usuários de mídias sociais em todo o mundo, gastando, individualmente, uma média de 2h27min todos os dias nas redes sociais. (We are social, 2022)

Já no relatório publicado em 2023, ao final da pesquisa, contou com a atualização dos dados e o aumento do alcance da conexão à Internet, o qual atingiu 64,4% das 8,01 bilhões de pessoas existentes no planeta. Da totalidade de habitantes, estima-se que 4,76 bilhões são usuários de mídias sociais em todo o mundo, gastando individualmente, em média, 1h56min todos os dias nestas plataformas, o que pode ser constatado que embora tenha ocorrido o aumento de usuários de mídias sociais, houve uma diminuição no tempo de consumo destas. (We are social, 2023)

Claro que mesmo ocorrendo essa redução de consumo, a inovação das mídias sociais e sua função pós-massiva continuam a trazer diversos malefícios inerentes ao uso exacerbado desses avanços tecnológicos, contribuindo para a disseminação de discursos que violam a dignidade do ser humano ou buscam atentar contra os estados democráticos de direito, dos quais seus interlocutores vêm se

utilizando do fundamento de que qualquer medida de controle ou moderação nas plataformas seria uma censura, isto é, uma violação ao direito à livre manifestação.

Mesmo havendo a garantia à liberdade de expressão, alguns discursos utilizados nessas plataformas acabam refletindo na violação de uma série de direitos humanos. Justo por isso, as mídias sociais vêm sendo pressionadas a investir em meios de controlar sua propagação de conteúdos impróprios através de mecanismos independentes e privados, uma vez que sanções aplicadas diretamente pelas plataformas, ou até mesmo o sistema judiciário, não estão conseguindo coibir de forma efetiva esse tipo de comportamento, devido a rapidez em que os conteúdos viralizam no meio digital.

Com isso, Mark Zuckerberg, CEO do Meta, criou um meio independente para analisar as sanções impostas aos usuários das plataformas *Facebook* e *Instagram* quando violassem os padrões da comunidade, surgindo, assim, o chamado *Oversight Board* (Comitê de Supervisão), como um comitê independente que julga casos relativos ao conteúdo de postagens e dos usuários dessas plataformas.

O comitê passou a julgar casos desde outubro de 2020, chamando a atenção para um caso emblemático envolvendo o ex-presidente norte-americano, Donald Trump, que incitou seus eleitores a invadir o Capitólio, em *Washington*. Após o incidente trágico ocorrido no dia 06/01/2021, o ex-presidente publicou em suas contas mensagens de apoio ao grupo invasor e incitando seu eleitorado contra os órgãos governamentais estadunidenses. Como consequência, teve suas contas no *Facebook* e *Instagram* suspensas por tempo indeterminado, por ter infringido os “Padrões da Comunidade” na categoria “indivíduos e organizações perigosas”, o que levou a plataforma recorrer ao Comitê para entender se a sanção aplicada estaria certa ou não.

A partir daí há uma preocupação acerca do direito dos usuários em exprimir suas opiniões e anseios nos meios eletrônicos de comunicação, como se manifestar sobre convicções e posicionamentos sem que extrapole os limites à liberdade de expressão. Se há limites no direito humano à liberdade de expressão quais seriam os elementos limitadores para a livre manifestação? A partir de quais padrões podem ser compreendidos e garantidos o direito humano à liberdade de expressão nas redes sociais?

É evidente o quanto a utilização das redes sociais consegue transcender os limites territoriais, resultando em impactos globalizados. O *Facebook* se tornou uma

rede extremamente suscetível para a manipulação de seus usuários, através de diversos mecanismos da inteligência artificial, tais quais os algoritmos, e vem servindo como ferramenta para aqueles que se utilizam de diversos discursos nocivos para alcançar – ou perpetuar – seus objetivos políticos e econômicos.

A Central de Transparência do *Facebook* faz uma análise trimestral sobre os conteúdos removidos nas plataformas do conglomerado, tendo, no trimestre de abril a julho/2021, sido removidos 6,2 milhões de conteúdos relacionados às organizações de ódio e 7,1 milhões relativas ao terrorismo, isso contando com a ajuda dos investimentos em inteligência artificial (IA) para detectar as postagens entre seus bilhões de usuários com contas ativas e nos mais variados idiomas espalhados pelo mundo, haja vista que 99,7% dos conteúdos de terrorismo que sofreram alguma sanção foram detectados por meio da IA, enquanto 0,3% foram denunciados por outros usuários, já em relação aos conteúdos de organização de ódio, 2,2% foram denunciados pelos usuários da plataforma e 97,2% foram identificados pela IA. (Facebook, 2022)

No entanto, mesmo com o auxílio da IA para identificar e remover o conteúdo das redes sociais, a ação não é suficiente para barrar discursos nocivos nas redes sociais ou retroagir seus efeitos negativos, uma vez que é necessário que haja uma abordagem pedagógica para orientar os usuários acerca dos limites à liberdade de expressão em suas manifestações *online*. E, nessa direção, é que a pesquisa acadêmica possui um papel de grande relevância para a sociedade, pois contribui para o reconhecimento dos problemas sociais, permitindo aos pesquisadores buscarem soluções para a resolução ou atenuação dessas situações.

Sendo assim, a justificativa desta pesquisa é buscar compreender o que é de fato liberdade de expressão e quais são os seus limites para que se possa exteriorizar opiniões, vontades, críticas, revoltas sem ofender o direito alheio ou incitar uma ruptura democrática, no meio virtual, por meio da análise de um caso específico, o qual se trata da sanção imposta contra ao ex-chefe de estado dos Estados Unidos.

Trata-se, portanto, de uma análise de caso, diante da repercussão do julgamento da remoção das contas do ex-presidente Donald Trump e a análise do referido caso pelo *Oversight Board*, tendo em vista o imenso impacto pelas postagens publicadas pelo então candidato à reeleição e o crescimento desmedido da polarização política aliado aos ataques de ódio e tentativas de atentados à democracia.

É, demasiadamente, temerário que figuras públicas e de cargos de alto escalões governamentais possam se utilizar das redes sociais, invocando o direito à liberdade de expressão como fundamental e absoluto, não havendo nenhuma sanção ou consequência de seus atos, mesmo que os efeitos dessas punições se atenham apenas às plataformas digitais.

Importante destacar nesse momento introdutório, que cada rede social possui suas regras e condições de uso, as quais são aceitas pelos usuários ao ingressar na plataforma e que estabelecem os conteúdos de postagens aceitos ou não, bem como suas sanções. Ocorre que há uma constante reivindicação de cerceamento ao direito à liberdade de expressão quando ocorrem punições nas plataformas, como a remoção do conteúdo impróprio até a suspensão ou exclusão de contas, como aconteceu com Trump.

As legislações internacionais, as quais são aplicadas nos julgamentos dos casos no Comitê de Supervisão, aparentam garantir o direito à liberdade de expressar opinião e transmiti-las por qualquer meio de comunicação, sem que haja interferência ou censura prévia para obstruí-las.

No entanto, mesmo que o direito à liberdade de expressão seja um direito fundamental positivado e reconhecido nas mais diversas diretrizes e convenções internacionais, ele não é e nem deve ser absoluto, sendo necessário entender seus limites e aplicá-los quando esses forem ultrapassados pelo interlocutor, bem como é necessário compreendê-los para que as próprias plataformas consigam elaborar suas políticas internas e regramentos, garantindo a proteção de direitos fundamentais de seus usuários.

E para compreender melhor a temática proposta, a problemática que se levantou foi de que maneira o *Oversight Board* pode garantir a proteção do direito humano à liberdade de expressão, a partir do estudo dos Padrões da Comunidade do *Facebook*, dos *standarts* Internacionais de direitos humanos sobre liberdade de expressão e a legislação norte-americano, tendo em conta a análise do caso referente à suspensão das contas do *Facebook* e *Instagram* do ex-presidente dos Estados Unidos, Donald Trump?

Trata-se, portanto, de uma pesquisa qualitativa, uma vez que tem por objetivo a compreensão do direito à liberdade de expressão os limites inerentes, preocupando-se “com aspectos da realidade que não podem ser quantificados, centrando-se na compreensão e explicação da dinâmica das relações sociais” , na

qual o método de abordagem empregado é o dedutivo, pois se utiliza de leis e teorias gerais para conseguir “explicar a ocorrência dos fenômenos particulares” (Silveira; Córdova, 2009, p.32; Diniz; Silva, 2008, p. 8). E através da aplicação desses entendimentos teóricos e legais pesquisados, será analisado no estudo a aplicação do direito humano à liberdade de expressão – premissa maior – nas manifestações ocorridas por meio das mídias sociais, tendo como caso a ser estudado o julgamento da suspensão das contas do Facebook e Instagram do ex-presidente Donald Trump – premissa menor.

O método de procedimento será o bibliográfico/monográfico, no qual consiste em um “levantamento de referências teóricas já analisadas, e publicadas por meios escritos e eletrônicos, como livros, artigos científicos, páginas de *web sites*” com o intuito de buscar uma solução ao problema proposto (Fonseca, 2002, p. 32).

A técnica aplicada é a revisão bibliográfica, tendo abrangido diversas bibliografias “já tornadas públicas em relação ao tema de estudo, desde publicações avulsas [...], revistas, livros, pesquisas, monografias teses” (Marconi; Lakatos, 2003, p. 182), dentre outros meios disponíveis, bem como documental.

Sendo assim, pode-se demonstrar que o presente projeto de pesquisa se encontra alinhado com a proposta “Direitos Humanos, Cidadania e Novos Direitos” do Programa de Pós-Graduação em Direito, da Universidade do Extremo Sul Catarinense – UNESC, uma vez que busca apresentar um estudo acerca da liberdade de expressão e seus limites, trazendo a análise de um caso concreto e de grande repercussão para a aplicação de medidas necessárias para conter o avanço do discursos atentatórios aos direitos humanos nas redes sociais, haja vista os impactos globais que a função pós-massiva das mídias sociais acarretaram na sociedade.

O projeto de pesquisa encontra-se, também, em consonância no que diz respeito à linha de pesquisa “Direitos Humanos, Novos Direitos e Litigiosidade”, do orientador, professor Dr. Gustavo Silveira Borges, tendo em vista que procura entender o impacto das novas tecnologias, como as das redes sociais, nos direitos humanos e na sociedade atual.

Por fim, foram realizadas pesquisas através do banco de dados do BDTD (Biblioteca Digital Brasileira de Teses e Dissertações) de trabalhos de dissertação e teses pelos termos, em conjunto, “liberdade de expressão” e “*oversight board*”, bem como “liberdade de expressão e “comitê de supervisão”, mas não foram encontrados nenhuma produção realizada até o momento. A mesma pesquisa foi realizada na

plataforma SciELO (*Scientific Eletronic Library Online*), com os mesmos termos da pesquisa acima citada, não tendo sido encontrado nenhum artigo científico publicado ainda.

À vista disso, a pesquisa que será realizada possui grande relevância, uma vez que é inédita nos estudos de mestrado e doutorado, bem como na publicação de artigos científicos sobre o tema proposto neste projeto.

Sendo assim, a presente pesquisa possui como objetivo geral estudar de que maneira o *Oversight Board* pode garantir a proteção do direito humano à liberdade de expressão nas mídias sociais, sendo dividido em três objetivos específicos, que são: descrever as revoluções da sociedade da informação até o surgimento das mídias sociais e a necessidade de reconhecimento dos novos direitos da Era Digital; compreender aplicação do direito à liberdade de expressão e identificar seus limites legais por meio dos padrões internacionais de direitos humanos, padrões da comunidade do *Meta* e a legislação norte-americana; e, por fim, analisar o julgamento da suspensão de contas do ex-presidente dos Estados Unidos, Donald Trump, através das premissas e conceitos estudados sobre liberdade de expressão, seus limites e o funcionamento do *Oversight Board*.

É o que passa a expor a seguir.

2 A SOCIEDADE CONTEMPORÂNEA, MÍDIAS SOCIAIS E A NOVA ÉGIDE DOS DIREITOS HUMANOS NA SOCIEDADE DA INFORMAÇÃO

A sociedade atual vem acompanhando diversas tecnologias e inovações serem incorporadas em seu cotidiano. Inteligência artificial, internet, do 3G a 5G, redes sociais, algoritmos, metaverso, chamadas virtuais em tempo real, *smart cities* são alguns exemplos desses avanços que surgiram nas últimas décadas. Entretanto, os usuários das tecnologias da informação e comunicação já estão tão habituados a dividir seu tempo e espaço que podem não vislumbrar todo o cenário que o mundo percorreu até chegar ao presente momento.

Analisar essa caminhada desde a concepção do que se entende pelo surgimento dos agrupamentos sociais, os quais começaram a formar as comunidades, as tecnologias que hoje parecem ser rudimentares, mas que deram início a todo o avanço industrial e tecnológico que está à disposição de cada indivíduo no planeta, compreender quando e como se deram esses eventos fazem com que se possa vislumbrar qual percurso a sociedade deve seguir e qual a direção que ela não deve ir.

De todo progresso alcançado advém ônus e bônus, inclusive os ocorridos na história da evolução industrial e tecnológica. Por isso que será descrito neste primeiro capítulo as revoluções da sociedade da informação até o surgimento das redes sociais, para que possa ser compreendida a necessidade do reconhecimento dos novos direitos na Era Digital e os meios para a proteção dos direitos humanos no mundo virtual.

2.1 A sociedade da informação e a vida *onlife*

A civilização humana passou por imensas mudanças desde sua concepção. A grande transformação que possibilitou que os seres humanos pudessem conviver em comunidade veio através da revolução agrícola, na qual o somatório das forças animais e humanas juntamente com a produção, transporte e comunicação resultou na urbanização e no surgimento das cidades. (Schwab, 2016, p. 18)

Após a revolução agrícola, o mundo se voltou para as Eras das Revoluções Industriais, que tiveram início nos séculos XVIII, com as ferrovias, a máquina a vapor

e produção mecânica, tendo sido denominada de Primeira Revolução Industrial. Com o surgimento da eletricidade, a criação de um modelo de linha de montagem e a consequente produção em massa, a Segunda Revolução Industrial começou a ser instaurada a partir do final do século XIX. A Terceira Revolução Industrial foi constituída “pelo desenvolvimento dos semicondutores, da computação em mainframe, computação pessoal e da internet” por meados dos anos 1960, tendo sido denominada de Revolução Digital ou do Computador. (Schwab, 2016, 18-19)

Essa nova Era da Revolução Digital pode ser abrangida, também, pela expressão “sociedade da informação”, uma sociedade constituída em um contexto pós-industrial e que se baseiam nas “transformações técnicas, organizacionais e administrativas que têm como ‘fator-chave’ não mais os insumos baratos de energia – como na sociedade industrial – mas os insumos baratos de informação propiciados pelos avanços tecnológicos na microeletrônica e telecomunicações”. (Werthein, 2000, p. 71)

Segundo Castells (2002, p. 39), essa evolução digital e de informação tiveram uma grande importância para a história, uma vez que “transformaram o cenário social da vida humana”. Essa transformação voltada às tecnologias da informação modificaram “a base material da sociedade em ritmo acelerado [...] apresentando uma nova forma de relação entre a economia, o Estado e a sociedade em um sistema de geometria variável”. E nesse contexto é que, justamente, a Terceira Revolução Industrial se vê pautada pela comunicação digital, tendo um alto crescimento de redes de interação de computadores, a partir das quais começaram a surgir novos formatos e “canais de comunicação, moldando a vida e, ao mesmo tempo, sendo moldadas por ela”. (Castells, 2002, p. 39-40)

Muito embora já existissem tecnologias informacionais com base na microeletrônica antes da Segunda Guerra Mundial, foi a partir desse período que as principais descobertas tecnológicas em eletrônica surgiram, muito por conta do imenso investimento bélico na área e, também, no campo da programação e da decodificação de armamentos e mensagens. Foi, então, que surgiu o primeiro computador programável e o transistor, o que se tornou o cerne da revolução da tecnologia da informação no século XX. (Castells, 1999, p. 76).

Todavia, foi somente após 1970 que as novas tecnologias se propagaram pelo mundo e contribuíram para a estimulação do desenvolvimento socioeconômico mundial, mas a tecnologia aplicada na época era irrelevante em comparação com as

tecnologias que existem hoje no mercado. Anos mais tarde o aprimoramento da tecnologia conseguiu chegar ao microcomputador que, associado ao roteamento de conexões, passou a revolucionar o mundo dos negócios. De início, o investimento era muito alto, o que fez com que essas tecnologias se tornassem pouco acessíveis naquele momento, entretanto, ao longo que a demanda e o interesse da população foram aumentando, essas novas tecnologias passaram a fazer parte do cotidiano da sociedade, cada vez com uma maior capacidade de funcionamento e processamento. (Castells, 1999, p. 76-81)

A Terceira Revolução Industrial é marcada, não apenas pelo advento dos computadores, mas também pela eclosão da internet, o que transformou as relações entre o ser humano e a máquina. Foi a partir daí que surgiu o conceito de globalismo, visto que a possibilidade de novas conexões entre indivíduos de diferentes localidades pode abreviar o espaço e intensificar a comunicação entre si. (Schwab, 2018, p. 38)

Toda essa nova estrutura fez com que a sociedade pós-moderna adquirisse uma nova característica: a das redes. Dessa forma, independentemente das diferenças geográficas, sociais, políticas, culturais ou qualquer outra característica limitadora, as redes e as mídias sociais formaram “um conjunto de nós interconectados”, os quais possuem uma infinita capacidade de expansão e integração de “novos nós” que conseguem se comunicar dentro dessa rede (Castells, 2002, p. 566).

Nesse sentido, Lévy (1999, p. 42-44) compreendeu que devido cada vez mais haver uma junção entre os espaços virtuais destinados ao trabalho com aqueles destinados à comunicação, fez com que a máquina – o computador – deixasse de ser o centro da tecnologia e passasse a ser apenas um elemento a mais para uma conexão ilimitada “cujo centro está em toda parte e a circunferência em lugar algum, um computador hipertextual, disperso, vivo, fervilhante, inacabado: o ciberespaço em si”.

Interpretar a sociedade atual - emergida da sociedade da informação e em rede - torna imprescindível compreender as transformações que estão ocorrendo nesse mundo globalizado e a reverberação de seus impactos, pois segundo Castells (1999, p. 44) a expressão “sociedade da informação” abrange uma modificação social, econômica e cultural da sociedade como um todo, o que resulta no surgimento de uma nova cultura baseada pela interação e hibridização das novas tecnologias com aquelas já preexistentes.

E não bastando os avanços que a sociedade da informação trouxe ao mundo e a forma como a informação passa a ser tratada, na Terceira Revolução Industrial¹, na virada do século XX, os avanços das tecnologias das redes de comunicação e informação trouxeram ao mundo “uma internet mais ubíqua e móvel, por sensores menores e mais poderosos que se tornaram mais baratos pela inteligência artificial e aprendizagem automática”, gerando uma ruptura da revolução digital iniciada no século anterior, surgindo assim uma Quarta Revolução Industrial. Essa nova revolução se traduziu em uma mudança significativa, na qual transformou sociedade do século XXI, uma vez que teve o poder de “fusão dessas tecnologias e a interação entre os domínios físicos, digitais e biológicos” se expandindo de forma mais célere e alcançando as mais diversas partes do mundo. (Schwab, 2016, p. 18-20)

A ascensão da nova Era Digital foi intitulada por Floridi (2014, p. 1-3) de *hiperstory*, na qual a sociedade se projeta a uma hiperconexão que vai além do território físico do usuário, tornando-se dependente das TICs. Além da internet mais ubíqua e a evolução da inteligência artificial, elemento essencial para a instauração da Quarta Revolução Industrial, Floridi adicionou a “Internet das coisas”, e a “*Web-augmented things*” como novos elementos dessa transformação. Toda essa relação humana ininterrupta com as TICs nada mais é que a experiência da vida *onlife* (Floridi, 2014).

Ao contrário do que Lévy trazia em seu entendimento acerca do ciberespaço, no qual havia um controle do usuário para entrar e sair da rede, isto é, havia a possibilidade de separar o espaço *online* do espaço *offline*, na vida *onlife* não há como diferenciar esses espaços, pois o usuário está inserido permanentemente nesses dois ambientes através de dispositivos como o GPS, *smartwatches*, *smartphones*, entre outros recursos (Floridi, 2014, p. 59-73).

Todas essas atividades executadas no modo *onlife* acabam resultando em um ambiente constituído de informações através dos fluxos de dados e isso possibilita com que as tecnologias da informação e comunicação consigam alterar a percepção de mundo de cada usuário, bem como têm a capacidade de criar novas realidades e promover uma interpretação informativa de cada elemento do mundo e da vida em

¹ Segundo Castells (1999, p. 43) a sociedade da informação, um dos produtos da Terceira Revolução Industrial, incorporou a informação como uma nova forma de capital e, também, de matéria-prima básica, tendo, de forma inédita superado os recursos naturais e a mão-de-obra.

que o indivíduo está inserido, esse ambiente informacional é chamado de infosfera (Floridi, 2014, p. 43).

Toda essa modificação de um ambiente dicotômico da informação, no qual havia uma separação entre o espaço *online* e o *offline*, para a infosfera, fez com que esse novo ambiente se tornasse mais sincronizado, deslocalizado e correlacionado, isto é, a normatividade e a territorialidade acabam se mostrando, de certa forma, mutável, uma vez que se passou a centralizar o espaço na relação entre os usuários, não mais o território em si (Floridi, 2014, p. 48; Floridi, 2020).

Entretanto, é importante pontuar que não somente das revoluções industriais que culminaram na eclosão e no desenvolvimento das tecnologias da informação e comunicação. Floridi (2014, p. 87), trouxe ao estudo da área uma nova visão sobre as revoluções da humanidade, abordando o viés das revoluções científicas, nas quais compreendeu que as três primeiras revoluções além de terem mudado a concepção do indivíduo de quem ele é, mas também sua autocompreensão.

Destaca-se, portanto, que a inserção das revoluções científicas, que serão abordadas a seguir, permitiram que a sociedade chegasse até a Quarta Revolução Industrial, não concomitantemente, mas de alguma forma se encontraram nesse novo *status* tecnológico do ciberespaço até a infosfera.

Sendo assim, a primeira revolução científica que pode ser pontuada foi a “revolução dos corpos celestes”, quando Nicolau Copérnico publicou seus estudos sobre o movimento dos planetas ao redor do sol. Seu tratado, sem sombra de dúvida, mudou definitivamente a autocompreensão que os seres humanos tinham de si, uma vez que se acreditava, naquela época, que a Terra era o centro do universo, “*nicely placed there by a creator God*”. A teoria do heliocentrismo causou uma imensa ruptura e reconsideração do universo, partindo daí que a palavra “revolução” começou a ser usada para denominar uma “transformação científica radical”. (Floridi, 2014, p. 87)

Depois da Revolução Copernicano, a segunda revolução se estabeleceu quando Charles Darwin apresentou sua tese sobre a origem das espécies pela seleção natural, isto é, a publicação de Darwin trazia a teoria de que “todas as espécies de vida evoluíram ao longo dos anos de ancestrais comuns por meio da seleção natural”, o que trouxe uma nova definição à palavra “evolução, transferindo o ser humano ao centro do “reino biológico” (Floridi, 2014, p. 88-89)

Segundo Floridi (2014, p. 89) as duas primeiras revoluções científicas impactaram a forma como a humanidade passou a se ver em relação ao universo e aos processos evolutivos, como se o controle da vida não estivesse mais nas mãos humanas, mas sim da própria natureza em si. Dessa forma, mesmo não estando mais ao centro desses processos, a sociedade ainda acreditava que, pelo menos, teria o controle total de seus pensamentos ou conteúdos mentais, uma vez que foi dessa forma que a humanidade passou a se enxergar como ser consciente a partir da célebre frase de René Descartes: “penso, logo existo”, isto é, pensava-se que a *“introspection was a sort of internal voyage of discovery of mental spaces [...] the mind was like a box: all you needed to do to know its contents was to look inside²”*.

Freud foi quem desmistificou o controle da mente humana pelo próprio indivíduo, segundo o psiquiatra “a mente humana também é inconsciente e sujeita a mecanismos de defesa”, ou seja, o indivíduo até pode construir narrativas, de forma consciente para defender seus atos e condutas, no entanto, muitas dessas ações são tomadas de forma inconsciente. E, dessa forma, a Terceira Revolução Científica se estabeleceu na descentralização do ser humano “do centro do reino da consciência”. (Floridi, 2014, p. 89-90)

Após as três primeiras revoluções científicas, as quais podem ser distinguidas, também, como uma revolução da autocompreensão humana, auxiliou, segundo Floridi (2014, p. 90), a humanidade a compreender que as significativas mudanças que foram experienciadas nas últimas décadas lhes auxiliou a chegar ao advento da Quarta Revolução, a qual pode ser conhecida pela Revolução da Informação.

Portanto, a Quarta Revolução Científica pode ser considerada como a revolução que deslocou o trabalho da mente humana, como os “únicos agentes inteligentes na infosfera”, ou seja o “o campo do raciocínio lógico, processamento de informações e comportamento inteligente” passou a ser feito também pelos dispositivos digitais, tendo a ciência da computação e as tecnologias da informação e comunicação realizado um papel de grande influência para que o ser humano pudesse revisar sua autocompreensão através desses novos meios conceituais, científicos e tecnológicos. (Floridi, 2014, p. 92-94)

² “[...] a introspecção era uma espécie de viagem interna de descoberta de espaços mentais [...] a mente era como uma caixa: para conhecer seu conteúdo bastava olhar para dentro”. (Floridi, 2014, p. 89, tradução livre)

Conforme já abordado aqui nessa pesquisa, o ambiente informacional se modificou de forma bastante insurgente a partir da virada do século XX e a Quarta Revolução, tanto a Industrial quanto a Científica acabaram se encontrando nessa dinâmica transformadora que impactou a vida da sociedade como um todo, ao remodelar a interação entre o ser humano e as TIC e isso é muito bem descrito por Floridi (2019, p. 94) no trecho colacionado abaixo:

I mentioned earlier that we are probably the last generation to experience a clear difference between online and offline environments. Some people already spend most of their time online. Some societies are already hyperhistorical. If home is where your data are, you probably already live on Google Earth and in the cloud. Artificial and hybrid multi-agents, i.e., partly artificial and partly human (consider, for example, a bank), already interact as digital agents with digital environments and, since they share the same nature, they can operate within them with much more freedom and control. We are increasingly delegating or outsourcing to artificial agents our memories, decisions, routine tasks, and other activities in ways that will be progressively integrated with us³.

Em suma, Floridi (2014, p. 96) demonstra em sua obra, na qual aborda sobre a Quarta Revolução, que os avanços tecnológicos que a sociedade está experimentando permite observar que as máquinas estão fazendo as mesmas funções que o ser humano fazia antes do seu surgimento. O autor pontua que determinadas funções exercidas por dispositivos digitais, tal como a inteligência artificial, retiraram, de certa forma, o que definia o ser humano como único:

It is humbling, because we share such a nature with some of the smartest of our own artefacts. Whatever defines us uniquely, it can no longer be playing chess, checking the spelling of a document or translating it into another language, calculating the orbit of a satellite, parking a car, or landing an aircraft better than some ICTs. You cannot beat ICT even at a random game such as rock-paper-scissors, because the robot is so fast that, in millisecond, it recognizes the shape that your hand is making, chooses the winning move, and completes it almost simultaneously. If you did not know better, you would think it was reading your mind.⁴

³ “Mencionei anteriormente que somos provavelmente a última geração a experimentar uma clara diferença entre ambientes online e offline. Algumas pessoas já passam a maior parte do tempo na vida. Algumas sociedades já são hiper-históricas. Se sua casa é onde estão seus dados, você provavelmente já mora no Google Earth e na nuvem. Multiagentes artificiais e híbridos, ou seja, parcialmente artificiais e parcialmente humanos (considere, por exemplo, um banco), já interagem como agentes digitais com ambientes digitais e, por compartilharem a mesma natureza, podem operar dentro deles com muito mais liberdade e controle. Estamos cada vez mais delegando ou terceirizando a agentes artificiais nossas memórias, decisões, tarefas rotineiras e outras atividades de maneiras que serão progressivamente integradas conosco.” (Floridi, 2019, p. 94, tradução livre)

⁴ “É humilhante, porque partilhamos essa natureza com alguns dos nossos artefatos mais inteligentes. O que quer que nos defina de forma única, já não pode ser jogar xadrez, verificar a ortografia de um documento ou traduzi-lo para outra língua, calcular a órbita de um satélite, estacionar um carro ou aterrizar um avião melhor do que algumas TIC. Não é possível vencer as TIC mesmo num jogo aleatório

Há, pelo menos, 8 anos Floridi escrevia sobre a transição que o mundo estava fazendo do ciberespaço e da sociedade em rede para a infosfera, ou seja, sobre o mundo *offline* sendo absorvido pelo *onlife world* e se preocupava acerca dos impactos que iriam recair sobre a Geração X, a qual, diferentemente das Gerações X e Y, não conseguirão reconhecer a diferença fundamental entre o mundo físico e o mundo digital, ficando cada vez mais “carente, excluída, deficiente ou pobre ao ponto de paralisia e trauma psicológico sempre estiver desconectada da infosfera”. (Floridi, 2014, p. 98)

E, por óbvio, Floridi não foi o único a se preocupar com os impactos, não somente positivos, que os avanços tecnológicos iriam causar à sociedade como um todo e essa é uma inquietação concernente à comunidade científica, na qual já observaram as transformações anteriores e suas repercussões negativas que ainda não foram totalmente solucionadas.

Schwab (2018, p. 38-39), por exemplo, defende que a comunidade mundial deve estabelecer três desafios para que as novas tecnologias, oriundas da Quarta Revolução Industrial, possam ser usufruídas por todos e de maneira saudável, os quais iriam consistir em:

O primeiro desafio é garantir que os benefícios da Quarta Revolução Industrial sejam distribuídos de forma justa [ao contrário do que ocorreu nas últimas revoluções, na qual a distribuição das riquezas e o direito ao bem-estar foram distribuídos e atribuídos de forma desigual]. [...] O segundo desafio é gerenciar as externalidades da Quarta Revolução Industrial no que diz respeito aos seus possíveis riscos e danos. Nas revoluções industriais anteriores foram feitos pouquíssimos esforços para proteger as populações vulneráveis, o ambiente natural e as gerações futuras contra o sofrimento que pode resultar de consequências não intencionais, dos custos da mudança, dos impactos de segunda ordem ou da má utilização deliberada das novas capacidades. [...] O terceiro desafio é garantir que a Quarta Revolução Industrial seja liderada por humanos e para humanos. Os valores humanos devem ser respeitados por si mesmos, em vez de ser apenas quantificados em termos financeiros. Além disso, estar centrado na humanidade é o mesmo que empoderar – e não ordenar – as pessoas como agentes com capacidade de realizar ações significativas.

Ainda para Schwab (2018, p. 40), o compartilhamento das benesses e um futuro projetado na centralização do ser humano não conseguirão ser solucionados

como pedra-papel-tesoura, porque o robô é tão rápido que, em milissegundos, reconhece a forma que a sua mão está a fazer, escolhe a jogada vencedora e completa-a quase simultaneamente. Se você não soubesse melhor, pensaria que ele estava lendo sua mente.” (Floridi, 2014, p. 96, tradução livre)

apenas por meio textos legislativos ou ações governamentais, mas sim em uma intervenção conjunta de instituições nacionais e internacionais, com a participação do mercado, movimentos sociais organizados e as *big techs*, buscando empoderar a sociedade para um melhor desenvolvimento tecnológico e econômico humanizado.

O que se pode observar nesse debate de ideias e consternações é que de um lado Floridi traz uma abordagem de preocupação intelectual dos impactos das tecnologias da informação e comunicação neste século, enquanto Schwab busca trazer à discussão a visão das consequências que as novas tecnologias estão trazendo e irão trazer ainda no desenvolvimento do mercado e da economia, o que, conseqüentemente, afeta a estrutura social ao redor do mundo, causando um enorme abismo de desigualdade na sociedade.

No entanto, nesse embate as indagações e incertezas de Schwab, aproximam de certa forma os dois autores, as incertezas originadas pelo rápido avanço tecnológico não conseguem ser elucidados com a implementação de novos sistemas ou planejamentos preventivos. O que deve ser trabalhado é a incorporação de uma “mentalidade que considere os efeitos no nível dos sistemas, o impacto sobre as pessoas, que continue a ser orientada para o futuro e esteja alinhada a valores comuns nos diversos grupos de partes interessadas”, as quais terão a tendência de absorver as novas tecnologias e equilibrá-las às necessidades da comunidade mundial com esses avanços. (Schwab, 2018, p. 43)

E tantas foram as revoluções industriais e tecnológicas, desde o surgimento do conceito de comunidade e agrupamentos sociais até os tempos atuais, os quais modificaram de forma considerável o âmago da sociedade, transformando sua estrutura e funcionamento. O estilo de vida não é mais o mesmo de dois séculos ou duas décadas atrás, novos conhecimentos trouxeram uma verdadeira gama de inovações na comunicação e, a partir da internet, surgiram as redes sociais para intensificar as interações sociais, mas que hoje já estão estabelecidas como meio de lazer e sedimentadas nas relações corporativas. O que será abordado no capítulo a seguir.

2.2 Do surgimento das mídias sociais à utilização das plataformas para disseminar atos de violação de direitos humanos

Pode-se considerar que as mídias sociais se tornaram um elemento central das tecnologias da informação e comunicação, tendo, inclusive, uma capacidade que foi além da centralização da comunicação em geral, uma vez que demonstrou possuir o poder de orientar as ações individuais e coletivas da sociedade.

Segundo o relatório da *Digital Media and Society*, as plataformas digitais possibilitaram uma conexão “do tipo um-para-um e um para-muitos”, fazendo com que os usuários dessas mídias pudessem fazer e manter contato independentemente do tempo ou da distância que estivessem, bem como aproximar pessoas que possuíssem opiniões semelhantes, transcendendo quaisquer “limites sociais, econômicos, culturais, políticos religiosos e ideológicos”. (Schwab, 2016, p. 100)

No início do século XXI, Castells (2002, p. 57) já previa que ao longo que a sociedade em rede iria se perpetrando na estrutura social resultaria na criação de uma variedade de comunidades virtuais, nas quais teriam a predisposição do viés social e político oriundos dos anos 90 e uma “preeminência da identidade como seu princípio organizacional.

Desde o início dos anos 2000, período em que o acesso à rede se dava por meio de máquina fixa em determinado local, até os tempos atuais, a forma de comunicação online mudou. O que antes centralizava a comunicação pelo computador, a mobilidade do espaço virtual trouxe para o seu cerne o usuário, que passou a ser a peça-chave dessa transformação.

Não somente o acesso à internet, mas as ferramentas para a comunicação na sociedade em rede aliada às tecnologias advindas da quarta revolução industrial fez com que as mídias sociais tomassem o gosto popular. O fácil acesso e a fácil comunicação entre os integrantes dessas comunidades virtuais tomaram o mundo e hoje já não há como imaginar um contexto sem que existam esses dispositivos.

As mídias sociais podem ter surgido, inicialmente, com uma tendência ao compartilhamento de situações da vida cotidiana de seus usuários, de comunicação instantânea entre indivíduos que estavam distantes territorialmente, mas suas funcionalidades se multiplicaram a ponto de fazerem parte da rotina de trabalho, do comércio e de lazer.

Inicialmente é interessante fazer um esclarecimento para a diferenciação entre o que se chama de mídias sociais e o que se conhece por rede sociais. Segundo Telles (2010, p. 7), embora exista uma utilização indistinta dos dois termos, as duas não significam a mesma coisa, na realidade as redes sociais são um elemento dentro

das mídias sociais, sendo que a primeira diz respeito às plataformas denominadas de “sites de relacionamento”, nas quais possuem um enfoque na reunião e interação de seus usuários na comunidade, com a exposição de perfil com dados pessoais compartilhados pelo próprio *user*, tais como fotos, vídeos, mensagens e outras funções.

Já as mídias sociais são todas e quaisquer plataformas que permitam uma “criação colaborativa de conteúdo, interação social e o compartilhamento de informações em diversos formatos”, ou seja, é o conjunto de todas as mídias existentes da Internet, podendo ser um microblog, compartilhamento de vídeos, compartilhamento de fotos, as redes sociais e outras plataformas. (Telles, 2010, p. 7-8)

As mídias sociais tiveram diversas modificações em seus formatos e funcionalidades desde o seu surgimento, por volta dos anos de 1980, tendo passado por quatro grandes transformações até chegar ao formato que conhecemos na atualidade. Sua origem se baseou em um formato de “fórum” para discussões em grupo, nos quais permitiam que os participantes criassem e trocassem conteúdo uns com os outros, essa foi chamada de “Primeira Era”. (Kaplan, 2016, p. 2)

A “Segunda Era” foi marcada com o início dos blogs na década de 1990, nos quais as pessoas relatavam publicamente os passos das suas vidas pessoais. Já nos anos 2000 foi, sem sombra de dúvidas, uma revolução em relação às TICs e com as redes sociais não foi diferente, com o estouro das pontocom (*dot-com*) as redes sociais voltaram a se evidenciar no plano virtual e trouxeram a “Terceira Era” das mídias sociais, foi nessa década que surgiram quatro mídias que permanecem até hoje como os principais sites do mundo: *Wikipedia*, *Facebook*, *YouTube* e *Twitter*. A “Quarta Era” deu uma nova configuração para as mídias assim que foram implementados novos formulários móveis, tirando a centralização do computador e as transportando para os dispositivos móveis, o que permitiu um maior acesso e compartilhamento em tempo real de postagens com informações referentes aos planos, momentos e localização de seus usuários. (Kaplan, 2016, p. 2-3)

As redes sociais se popularizaram com o advento da vida *onlife*, todos querem ficar conectados, mostrar seu dia a dia, vender seus produtos e serviços, assistir filmes, séries, fazer críticas sociais, mostrar seus momentos de lazer ou buscar esses momentos dentro das plataformas.

E para compreender melhor o funcionamento e estrutura das redes sociais, Recuero (2009, p. 25) pontuou que essas plataformas possuem dois elementos essenciais: os atores e as conexões. O primeiro elemento, que são os atores, diz respeito aos indivíduos participantes da rede social, os quais “atuam de forma a moldar as estruturas sociais, através da interação e da constituição de laços sociais”.

A autora, ainda, constata que em razão da distância entre os usuários, em um primeiro momento os atores podem não ser identificados nesses planos, mas que através das “representações dos atores sociais, ou com construções identitárias do ciberespaço”, eles se traduzem por meio de “um weblog, por um fotolog, por um twitter ou mesmo por um perfil do Orkut”. (Recuero, 2009, p. 25)

Óbvio que diante da obra apresentada por Recuero ser datada no ano de 2009, seus exemplos de redes sociais estão, de certa forma, defasados, visto que fotologs e o Orkut não existem mais, tendo sido substituídos por outros tipos de redes sociais mais modernas. Hoje há uma diversificação muito maior que no início dos anos 2000, além de que os meios de comunicação tiveram diversas modificações.

Na atualidade a fácil aderência na utilização das redes sociais na sociedade, muito por causa da grande abrangência na conexão em nível global, trouxe uma imensa variedade de novas plataformas, hoje há um total de 5,16 bilhões de usuários ativos nas mídias sociais em geral, segundo o relatório *Digital 2023 (We are social, 2023, p. 10)*, publicado pela *We are social* em parceria com a *Meltwater*.

Ademais, entre as principais redes sociais com mais números de usuários, se encontram: o *Facebook*, na primeira colocação com 2,9 bilhões de usuários ativos; o *Instagram*, está na quarta colocação com 2 bilhões; o *Tiktok*, na sexta posição com 1 bilhão; *Snapchat*, em décimo lugar com 635 milhões; e o *Twitter*, em décimo quarto lugar com 556 milhões de usuários ativos. (We are social, 2023)

Entretanto, mesmo que as redes sociais tenham se modificado nas últimas duas décadas, a representação dos atores sociais dentro das plataformas segue a mesma configuração exposta por Recuero (2009, p. 25), na qual são definidas como:

[...] espaços de interação, lugares de fala, construídos pelos atores de forma a expressar elementos de sua personalidade ou individualidade. Assim, um primeiro aspecto relevante para este estudo é a característica da expressão pessoal ou pessoalizada na Internet. [...] Essas apropriações funcionam como uma presença do “eu” no ciberespaço, um espaço privado e, ao mesmo tempo, público. Essa individualização dessa expressão, de alguém “que fala” através desse espaço é que permite que as redes sociais sejam expressas na Internet. (Recuero, 2009, p. 25-27)

Dentre a construção dos espaços sociais pelos atores nas redes sociais, temos como segundo elemento essencial dessas plataformas: a conexão, as quais são geradas por meio das “percepções que são construídas pelos atores”. (Recuero, 2009, p. 27)

Assim, entender como os atores constroem esses espaços de expressão é também essencial para compreender como as conexões são estabelecidas. É através dessas percepções, que são construídas pelos atores, que padrões de conexões são gerados, ou seja, é a partir delas que serão estabelecidas as estruturas de cada plataforma. Portanto, as conexões são, nada mais nada menos, que os laços sociais construídos dentro das redes sociais através da interação entre os usuários. (Recuero, 2009, p. 27-30)

Essas interações nas redes sociais são chamadas de assíncronas e síncrona, na qual a primeira diz respeito às comunicações que não possuem uma expectativa de resposta imediata e isso podemos ter como exemplo os e-mails e fóruns de discussões - muito embora não sejam mais tão populares. Já as interações síncronas se referem aquelas que ocorrem em tempo real, nas quais existe uma expectativa de resposta/interação imediata entre seus usuários, o que na atualidade ocorre com grande parte das redes sociais. (Recuero, 2009, p. 31-32)

À vista do que foi exposto até o momento, pode-se perceber que desde que as redes sociais surgiram, muitas foram as transformações que passaram até chegar nas plataformas que são usadas cotidianamente pela sociedade global. Entretanto, por mais que diversos benefícios tenham surgido com esses fenômenos tecnológicos, as redes sociais também serviram como instrumento para efetivação de atos de violação de direitos humanos.

As redes sociais, desde sua concepção, sempre contaram com um dinamismo nas suas interações. No entanto, o que no início da sociedade em rede se tinha uma comunicação dividida entre o ambiente online e offline, isto é, só podia ter acesso ao que acontecia no mundo virtual quando o usuário se conectava a internet, agora temos uma interação síncrona com o mundo online em grande parte das plataformas, o que já foi exposto nesta pesquisa como a chamada “vida *onlife*”.

E conforme já explanado, também, os avanços das tecnologias das redes de comunicação e informação trouxeram ao mundo “uma internet mais ubíqua e móvel, por sensores menores e mais poderosos que se tornaram mais baratos pela

inteligência artificial e aprendizagem automática”, que se traduziu em uma mudança significativa no comportamento do mundo virtual no século XXI, a partir da “fusão dessas tecnologias e a interação entre os domínios físicos, digitais e biológicos”, os quais se expandiram de forma mais célere e alcançando as mais diversas partes do mundo. (Schwab, 2016, p. 18-20)

A partir daí, as redes sociais passaram a ter um papel de extrema relevância na vida cotidiana de cada indivíduo, pois não apenas servia para facilitar a conexão e o relacionamento entre qualquer pessoa de qualquer lugar no planeta, como passou a ter um poder na abrangência de emissão e na formação de opiniões. Devido à proporção dos avanços tecnológicos e a hiperconectividade, conteúdos envolvendo violações de direitos humanos se beneficiam da função pós-massiva das mídias sociais, permitindo um alcance cada vez mais rápido e em escala mundial.

No entanto, as redes sociais não foram criadas com a intenção de serem utilizada para disseminação do discurso de ódio, de conteúdos envolvendo *fake news* ou de qualquer outro assunto que atente contra dos direitos humanos, entretanto, os “engenheiros do caos”⁵ se deram conta de que essas modernas ferramentas de comunicação eram um poderoso meio para difundir e provocar “fortes emoções, polêmicas, indignação e raiva” e que, conseqüentemente, essas comoções passavam a engajar os usuários que compartilhavam dos mesmos sentimentos dentro dessas plataformas. (Empoli, 2019, p. 46)

O *Facebook* é um exemplo de como uma rede social pode ser um extraordinário mecanismo para disseminação desse fenômenos da era da pós-verdade, transcendendo as dimensões da plataforma e passando a fazer parte da realidade de seus usuários, como no caso da campanha presidencial de Donald Trump, em 2016, que se utilizou dos dados colhidos e fornecidos pela *Cambridge Analytica*⁶ para impulsionar aos seus eleitores conteúdos de discursos de ódio, tendo

⁵ Empoli descreve os “Engenheiros do Caos” como aqueles que perceberam que as redes sociais poderiam ser grandes aliadas na disseminação de ideais duvidosos e antidemocráticos, utilizando-se das ferramentas existentes nessas plataformas para manipular os usuários através de *fake news*, discurso de ódio, dentre outros mecanismos.

⁶ O escândalo de vazamento de dados do *Facebook* pela *Cambridge Analytica* ocorreu quando a empresa de consultoria utilizou aplicativo de enquete na plataforma, o *thisisyourdigitallife*, e colheu, além dos dados dos usuários que fizeram o teste, os de seus amigos na rede social. Segundo a reportagem, 270 mil pessoas fizeram a enquete, mas foram coletados os dados de 50 milhões de usuários. Essa forma de coleta de dados possui o propósito de traçar perfis de personalidade dos usuários das redes sociais, para criar nichos de acordo com a qualidade encontrada na enquete e, enfim, direcionar a exibição de anúncios específico de acordo com seus perfis e, claro, tendo como escopo alcançar certos objetivos políticos (BBC News, 2018; Demartini, 2018).

percebido que “a indignação, o medo, o preconceito, o insulto, a polêmica racista ou de gênero se propagam nas telas e proporcionam muito mais atenção e engajamento que os debates enfadonhos da velha política”, o mesmo aconteceu com a campanha para o referendo do *Brexit*, no Reino Unido, em 2016, e a campanha presidencial de Jair Bolsonaro, no Brasil em 2018 (Empoli, 2019, p. 52).

A extração de dados dos usuários das plataformas digitais permitiu que narrativas que possuíssem o intuito de manipular pudessem se disseminar de forma mais fácil, justo porque essas agências publicitárias conseguem direcionar conteúdos personificado para o perfil de cada usuário através da tecnologia persuasiva⁷ e do *growth hacking*, que funcionam por meio dos algoritmos, os quais passam a mostrar posições e opiniões similares entre membros de um mesmo grupo social ou ideológico. Essa técnica faz com que o usuário acredite que a comunidade esteja pensando e agindo do mesmo modo que ele e esse mecanismo, invisível aos olhos dessas pessoas, cria bolhas de manipulação que se direcionam às polarizações. Assim, estabeleceu-se um sistema inclinado a conteúdos tais como as *fake news* e discurso de ódio, haja vista que por ter um engajamento maior dentro da comunidade nessas publicações, as empresas acabam gerando cada vez mais lucros para si mesmas. (O dilema..., 2020)

Schwab (2016, p. 101) já havia identificado que a utilização das redes sociais para a manipulação era um problema em ascensão, tendo em vista que essa dinâmica de compartilhamento de conteúdo tinha um imenso potencial de direcionar “a tomada de decisões e causar riscos para a sociedade civil”, fazendo com que “tudo que lemos, compartilhamos e vemos nas mídias sociais moldam nossas decisões políticas e cívicas”, podendo essas ferramentas serem utilizadas por agentes extremistas para manipular determinados grupos em favor de suas causas.

E justo por isso, é que a utilização das ferramentas de persuasão e manipulação que as redes sociais oferecem de forma democrática somadas aos conteúdos ofensivos, que gradativamente mais, líderes extremistas estão se beneficiando das plataformas digitais para alcançar – ou até perpetuar – seus

⁷ A tecnologia persuasiva pode ser identificada como uma “interseção entre a área de design e a tecnologia. Basicamente, trata-se do desenvolvimento de produtos capazes de influenciar atitudes e modificar o comportamento de seus usuários”. Esse mecanismo foi criado, intencionalmente, para persuadir e influenciar a sociedade, por meio de manipulação de processos psicológicos básicos da cognição social (Joaquim, 2021)

objetivos políticos e econômicos, fazendo com que isso impacte de forma exponencial na vida da comunidade, seja ela local ou a nível global.

Para Subirats (2011) há quem acredite que desde a primeira geração da midiaticização da política – através do rádio e da televisão – afastou o debate político do cidadão o transformando em algo mais virtual. A segunda geração das mídias – pelas redes sociais – permitiu que o controle ilimitado de dados somado a um sofisticado *marketing* político configurasse uma grande probabilidade de manipulação informativa. Por isso, não se pode ter a ingenuidade do grande peso que as instituições públicas e o mercado possuem para criar novas estruturas hierárquicas, de controle e monopólios nessa sociedade em rede.

Contudo, não há como ignorar que essa nova era informacional e de comunicação implicou em um crescente sentimento cívico e associativo, através de comunidades virtuais, os quais aumentaram o debate político e reforçaram novas autonomias sociais, o que traz um ponto positivo para a utilização das redes sociais pela população como forma de reacender o dever cívico do cidadão. Movimentos como o 15M e Primavera Árabe tiveram reflexos ao redor do mundo e mobilizaram pautas similares em Israel, Índia e Nova York, o que traz a reflexão de que as redes sociais podem servir como uma grande aliada no recrutamento e no engajamento da população na busca de soluções de problemas regionais ou globais. (Subirats, 2011)

Na contramão desse cenário positivo, descrito por Joan Subirats, até porque se for levada em consideração o contexto mundial na data de sua obra, as redes sociais estavam, de fato, sendo grandes aliadas dos movimentos sociais, auxiliando no crescimento desse dever cívico-democrático do cidadão em buscar engajamento nos problemas existentes na sociedade.

Todavia, D'ancona (2018, p. 21) apontou, posteriormente, que a era da pós-verdade gerou um grande impacto na forma em que a população passou a se portar diante dos fatos apresentados a ela. O que se imaginava era que “os eleitores reuniam fatos, tiravam conclusões desses fatos e assumiam posições a respeito das questões com base em suas conclusões e escolhiam um partido político de forma correspondente”. Ocorre que na prática os cidadãos estavam escolhendo seus partidos “com base em afiliações de valor, adotavam as opiniões da tribo, desenvolviam argumentos para apoiar essas opiniões e (só então) selecionavam os fatos para reforçar essas alegações”.

Segundo Empoli (2019, p. 46-47) as manipulações políticas propagadas através de teorias conspiratórias, *fake news* e discurso de ódio tiveram sucesso nas mídias sociais porque elas “provocam fortes emoções, polêmicas, indignação e raiva”, fazendo com que o engajamento nos cliques dessas postagens oportuniza um usuário cada vez mais dependente dessa rede.

Um estudo realizado pelo Instituto de Tecnologia de Massachusetts (MIT), em 2018 e publicado na revista *Science*, comprovou que uma notícia fraudulenta possui, em média, 70% mais chances de ser compartilhada nos meios digitais, atingindo seus 1.500 usuários seis vezes mais rápido do que uma notícia verdadeira, além disso, a pesquisa mostrou que o compartilhando desse tipo de conteúdo gerado por políticos engajam mais rápido porque “*a norm-breaking candidate who tells lies appeals to aggrieved constituencies because those voters see norms as illegitimately imposed by the establishment*”⁸. (Brown, 2020)

O sentimento de insatisfação da população contribuiu para a crise da democracia representativa e, conseqüentemente, do processo democrático eleitoral, trazendo um crescimento de movimentos (ultra)nacionalistas. O *Facebook* auxiliou esse movimento, pois suas ferramentas de compartilhamento de conteúdo e engajamento “funcionou como um formidável multiplicador, nutrindo-se dos ingredientes mais díspares para disseminar uma epidemia de cólera que se transferiu da dimensão virtual para a realidade”. (Empoli, 2019, p. 49-97)

Trump é um excelente exemplo de como as redes sociais alavancou sua campanha e o levou à Casa Branca em 2016. O eleitor norte-americano tinha total ciência dos desvios de caráter do então candidato à corrida presidencial, mas o sentimento que ele desencadeava em seu eleitorado de “fúria, impaciência e atribuição de culpa”, fazia com que fosse irrelevante se ele estivesse falando a verdade o não, porque isso causava um sentimento de que se tratava de um candidato diferenciado. (D’Ancona, 2018, p. 37)

Entretanto, todo esse carisma populista que havia em torno de Trump não seria suficiente sem que houvesse um agente externo para consolidar seu caminho rumo à presidência dos Estados Unidos. E foi nesse cenário que as redes sociais tiveram uma importante utilidade ao usar de dados coletados pela *Cambridge*

⁸ “um candidato que quebra as normas e que conta mentiras apela aos círculos eleitorais prejudicados porque esses eleitores veem as normas como impostas ilegítimamente pelo sistema”. (Brown, 2020, tradução livre)

Analytica, em enquetes no *Facebook*, para direcionar conteúdos específicos – envolvendo *fake News*, discurso de ódio e teorias conspiratórias – para potenciais eleitores. (Empoli, 2018, p. 53)

O mesmo ocorreu na campanha da retirada do Reino Unido da União Europeia, o chamado *Brexit*, os estrategistas da campanha se deram conta que a população gerava muito mais engajamento quando incluíam algum elemento que causasse grande emoção ou comoção, passando de centenas para milhares de compartilhamentos. No Brasil, o então candidato à presidência da república, Jair Bolsonaro, utilizou-se tanto do *Facebook* quanto do aplicativo de *WhatsApp* para disseminar conteúdos fraudulentos e de discurso de ódio. Para Empoli, os engenheiros do caos se deram conta que para um conteúdo viralizar e conseguir manipular um determinado nicho de eleitores, é preciso que o conteúdo sensibilize o usuário por meio de um sentimento – já existente – “indignação, medo, preconceito, insulto, polêmica racista ou de gênero”, ficando cada vez mais distante dos debates arcaicos da “velha política”. (Empoli, 2019, p. 53)

É notória a transformação que a civilização humana passou no último século, tendo tido um destaque para a virada do século XX para o XXI, possibilitando um avanço tecnológico um pouco diferente daqueles projetados em filmes e histórias de ficção.

Todavia, o impacto que a vida *onlife* trouxe no cotidiano atual e as ferramentas utilizadas nesse novo modo de viver possuem um grande poder de influência e determinação sobre o ser humano, no qual é de extrema relevância que a comunidade acadêmica possa identificar as repercussões negativas e traçar soluções para que essas tecnologias não resultem em novas violações de direitos humanos.

2.3 Da necessidade de novas configurações para a proteção dos direitos humanos na Era Digital

As últimas décadas deixaram a sociedade diante de uma insegurança jurídica devido ao crescimento exponencial das TICs e a escassa regulamentação sobre a temática (ou a falta dela). Essa lacuna jurídica, seja ela refletida nos ordenamentos jurídicos pátrios ou na legislação internacional, vêm causando repercussões negativas aos direitos humanos desde que a vida *onlife* se impôs no

cotidiano mundial. A comunidade acadêmica tem buscado encontrar soluções para a resolução desses conflitos, principalmente aqueles decorrentes da utilização das redes sociais para disseminação de conteúdo de ódio e desinformação, os quais estão transcendendo os limites territoriais, resultando em impactos globalizados.

Os direitos humanos perpassaram séculos garantindo direitos essenciais para assegurar a dignidade humana da sociedade, tendo sido conquistados os direitos de primeira dimensão, por volta dos séculos XVIII e XIX, quando se positivaram os direitos civis e políticos, celebrados pelas declarações de direitos de Virgínia (1776), da França (1789), Constituição Americana (1787), Constituições Francesas (1791 e 1793) e Código Napoleônico (1804) (Wolkmer, 2002, p. 13-14).

Os direitos de segunda dimensão conquistaram os direitos sociais, econômicos e culturais, os quais se fundaram nos princípios da igualdade, destacando-se as Constituição Mexicana (1917), Constituição Alemã de Weimar (1919), Constituição Espanhola (1931) e o Texto Constitucional do Brasil, em 1934. Já os direitos de terceira dimensão trouxeram a garantia dos chamados direitos metaindividuais, relativos aos direitos coletivos e difusos e os direitos de solidariedade, no Brasil teve destaque a Constituição Federal de 1988, o Estatuto da Criança e do Adolescente (1990) e o Código de Defesa do Consumidor (1990). O final do século XX apresentou uma nova gama de direitos na quarta dimensão, assegurando direitos à biotecnologia, à bioética e à regulação da engenharia genética, traduzidas pelo Código de Nuremberg (1947), pela Declaração de Helsinque (1964), a Lei Brasileira da Biossegurança (1995) e a Lei de Doação de Órgãos (1997). (Wolkmer, 2002, p. 14-21)

A quinta dimensão se refletiu nos direitos oriundos “das tecnologias de informação, da Internet, do ciberespaço e da realidade virtual em geral”, uma vez que a chegada no Novo Milênio, nos anos 2000, impactou numa mudança significativa de uma sociedade industrial para uma “sociedade da era virtual”. O impacto no direito, na sociedade e nos bens culturais que a evolução das tecnologias de informação e comunicação receberam através “das redes de computadores, do comércio eletrônico, das possibilidades da inteligência artificial e da vertiginosa difusão da internet” foi de ordem global. (Wolkmer, 2002, p. 21-22)

Sendo assim, a partir do surgimento de novas necessidades sociais, assim como denomina Wolkmer (2002, p. 10), foram a constituição de novas categorias jurídicas, distintas daquelas tradicionais. Os novos direitos não conseguem ser

combatidos pela “dogmática jurídica clássica”, uma vez que elas possuem um caráter relativo, não podendo ser classificadas como direitos difusos ou metaindividuais, pois carecem de uma imposição de “instrumentos jurídicos adequados para viabilizar sua materialização e para garantir sua tutela jurisdicional”.

É evidente, conforme já abordado até o momento, que a última revolução digital teve grande repercussão na vida cotidiana da civilização humana, bem como o alcance infinito da Internet e a utilização das redes sociais para manipulação de seus usuários, ocasionaram novos conflitos jurídicos e violações aos direitos humanos.

Justo por isso, é demasiadamente necessário que se busque instaurar uma nova regulamentação que consiga “controlar e proteger os provedores e os usuários dos meios de comunicação eletrônica de massa”, uma legislação que condicione a esses novos direitos uma qualidade de vida, bem-estar e materialidade social. (Wolkmer, 2002, p. 22-26)

Na conjuntura nacional, como pode ser constatado, o Brasil já conquistou a positivação de alguns direitos emergentes, tais quais: o Marco Civil da Internet (2014), direito ao esquecimento; a tipificação dos cibercrimes; a interceptação de dados telemáticos; o teletrabalho; a Lei Geral de Proteção de Dados (2018), dentre tantos outros. (Takano; Silva, 2020, p. 10-11)

Entretanto, conforme denunciaram Takano e Silva (2020, p. 11), para que esses novos direitos consigam ser reconhecidos e positivados, seja através do direito internacional ou dos ordenamentos pátrios, esbarram em obstáculos que criam lacunas legislativas e “provocam instabilidade social e insegurança jurídica”, haja que vista que a imposição da norma jurídica é obsoleta e morosa demais para a rapidez com que as tecnologias da informação e comunicação evoluem a cada dia, sendo impossível seu acompanhamento e a atualização da norma defasada:

O ciberespaço precisa da neutralidade para garantir o respeito aos princípios dos direitos humanos universais. Para tanto, os direitos provenientes do novo paradigma tecnológico necessitam de reconhecimento e concretização para que haja um crescimento tecnológico saudável e efetivo sem restrições e privações à dignidade da pessoa humana. (Takano; Silva, 2020, p. 12)

Entretanto, ainda existem obstáculos a serem superados para que se possa aplicar uma nova ordem jurídica para abranger esses novos direitos, uma vez que, segundo Zibetti, Pilau Sobrinho e Miranda (2018, p. 114-116) ainda temos presente no direito internacional uma “lógica jurídica centrada na figura dos Estados e

estruturada a partir da noção de soberania”, as chamadas fontes clássicas, as quais mesmo com a repercussão do crescimento de tratados internacionais a partir do século XIX conseguiu se esquivar dessa lógica clássica, não resultando em impactos significativos a sociedade internacional, até porque se vislumbrou a flagrante “incapacidade dos estados para lidarem sozinhos com problemas de amplitude transnacionais”.

Após o cenário de alta tensão resultante da Segunda Guerra Mundial, a criação de um órgão de cooperação internacional, a ONU, teve um papel importante para amenizar os problemas oriundos com o fim do regime nazista, o que “numa conjuntura em que a guerra definitivamente deixava de ser uma saída viável às oposições políticas, as organizações internacionais aos poucos passavam a ser vistas sob outra perspectiva”, causando uma ruptura nesse sistema internacional anteriormente baseado na soberania dos Estados, formalizando a distinção entre a sociedade internacional e os Estados. (Zibetti; Pilau Sobrinho; Miranda, 2018, p. 116)

Consoante destacado por Zibetti, Pilau Sobrinho e Miranda (2018, p. 115) a configuração para a abertura de um futuro direito internacional pode ter se dado em um contexto de revolução industrial, uma vez que:

segundo uma dinâmica já visível no longínquo advento da Carta Magna Inglesa e na proclamação do *Bill of Rights*, consolidou-se a tendência que, no futuro, levaria à emergência do indivíduo como sujeito de direito internacional. No caso da primeira, o avanço tecnológico, que a partir de então não conheceria mais freio, reduziu as distâncias intensificando as relações internacionais, mas também os problemas delas advindos. (Zibetti; Pilau Sobrinho; Miranda, 2018, p. 115)

Sendo assim, desde o arcabouço dessa “sociedade internacional” o direito se habituou entre dois sistemas que são divididos entre “o antigo ainda operacional e dotado de suas fontes tradicionais, e o novo, ainda não totalmente operante e presente em alguns segmentos específicos”, entretanto, mesmo com o avanço do direito internacional para buscar uma adaptação aos novos direitos, sua forma clássica de resolução de problemas não conseguem acompanhar “os desafios dessa nova sociedade globalizada cada vez mais carente de respostas rápidas para problemas urgentes”. (Zibetti; Pilau Sobrinho; Miranda, 2018, p. 118)

É notório o impacto que as revoluções da sociedade pós-moderna tiveram na formalidade do direito como um todo, bem como a globalização dos problemas oriundos dessas evoluções tornaram o direito clássico ineficiente demais para acompanhar e oferecer elucidaciones eficazes às complexidades cotidianas globais.

Devendo ser buscadas alternativas que de “realinhamento do direito em relação às demandas da sociedade internacional globalizada, promovendo a sinergia entre os diferentes instrumentos de organização social”. (Zibetti; Pilau Sobrinho; Miranda, 2018, p. 122)

Urge, então, uma nova abordagem do direito constitucional (ou neoconstitucionalismo), no qual precisa identificar todo o somatório de transformações ocorridas na sociedade e no mundo jurídico, principalmente aquelas transformações consolidadas desde o final do século XX, podendo ser estabelecido configuração jurídica: o constitucionalismo digital. (Barroso, 2005, p. 15; Takano; Silva, 2020, p. 8)

As revoluções tecnológicas da informação e comunicação, iniciadas na virada do século XX para o século XXI, trouxeram novas características para a sociedade e ferramentas de comunicação, dentre elas as redes sociais. Ao longo que esses *sites* de relacionamento foram se aprimorando, permitindo um compartilhamento de dados em tempo real e ilimitado territorialmente, fez com que seus usuários se sentissem mais à vontade para manifestar suas opiniões, crenças, vontades, incluindo nesse *rol* manifestações nocivas que incitam ou apoiam a violação de direitos humanos fundamentais.

Esse ideal libertário foi compartilhado por John Perry Barlow, um dos fundadores do *Electronic Frontier Foundation* (EFF), o qual acreditava que a própria comunidade da internet fosse capaz de criar suas próprias regras de forma justa e democrática, sem que houvesse a interferência de leis de qualquer Estado-nação, ou seja, a internet não poderia nem deveria se sujeitar a aplicação de regras genéricas impostas pelos *outsiders* do ciberespaço, pois a “tirania da maioria” forçaria a minoria a se submeterem a regras que não iriam aceitar. (Suzor, 2019, p. 88-89)

Segundo Suzor (2019, p. 89) a nova democracia do ciberespaço buscava promover um ideal libertário de que caso o usuário não aceitasse as regras dadas por determinada comunidade, teria a liberdade de deixá-la e construir uma nova comunidade através de ideais individuais que acreditasse ser pertinente, uma vez que a imposição, pelo Governo, de censura e vigilância na infraestrutura da internet, na qual interferisse em seu ideal de neutralidade originária, era o que os libertários do ciberespaço sempre temeram.

Ocorre que essa visão utópica libertária não poderia prever que o espaço virtual fosse ser, futuramente, dominado por pequenos grupos, os quais detêm o poder de controle dos meios virtuais. Hoje, a saída de um membro do ciberespaço, por não

concordar com suas regras, não gerará nenhum impacto efetivo para melhora de seus ordenamentos internos, muito mais fácil que leis, construídas democraticamente e fora do ambiente virtual, busquem se adaptar de forma mais efetiva às necessidades dos usuários. (Suzor, 2019, p. 89)

Ainda segundo o autor, os pequenos grupos que controlam os maiores meios digitais não são as entidades mais poderosas do mundo, no entanto, são os atores-chave do controle de nossas vidas, uma vez que toda essa centralidade do controle das mídias sociais os torna suscetíveis a influência de tantos outros influentes que possuem o desejo de regular nosso comportamento para alcançar seus próprios objetivos. (Suzor, 2019, p. 90)

E, portanto, toda essa influência que as *big techs* passaram a ter sobre seus usuários se traduziram por meio dos “termos de uso” para o acesso às mídias sociais em geral. Pode-se perceber que quando o usuário resolve ingressar em determinada plataforma, ele deve, antes de tudo, concordar com as regras estabelecidas prévia e unilateralmente, não tendo autoridade alguma para alterar qualquer parte que não concorde.

As mídias sociais passam, tal qual o *Facebook*, a exigir do usuário a concordância de “termos de uso” em seu ingresso na plataforma, no qual estariam sujeitos a esse regulamento interno em que a plataforma teria o total controle para determinar seu funcionamento e como sua comunidade pode e deve agir. A ideia da imposição de termos para utilização das redes sociais foi criada com o intuito de isentar as empresas de responsabilidades provenientes de violações cometidas por seus usuários, baseando-se que as redes sociais são consideradas espaços privados. (Suzor, 2018, p. 3).

Suzor (2018, p. 3) pontuou que as plataformas digitais possuem, de fato, a faculdade e o direito de regular como seus sistemas poderão ser utilizados, pressupondo-se que seus usuários são atores autônomos e racionais em um contexto liberal. No entanto, cada vez mais está evidente o papel que essas plataformas estão exercendo na autonomia da vida cotidiana de seus usuários. E isso pode ser visualizado em escândalos como o da *Cambridge Analytica* e a influência direta aos usuários por meio da manipulação algorítmica, seja para consumo ou incitação de polarizações políticas e ideológicas. Carece, portanto, um ordenamento jurídico que detenha clareza no conceito de responsabilidade das mídias sociais, uma vez que “*there is no consensus about the extent to which users need protection from the*

*governance decisions of platforms themselves*⁹, tornando-se difícil a discussão sobre a legitimidade das preocupações de limites dos usuários e das próprias plataformas. (Suzor, 2018, p. 4)

A vida *onlife* que a sociedade global está exposta atualmente não consegue mais se desvincular do mundo virtual, pois é onipresente e já causa impactos – sejam eles positivos ou negativos – mesmo que o indivíduo não queira. A função pós-massiva das redes sociais amplifica esses impactos e os tornam cada vez mais globalizados, atingindo os mais diversos grupos sociais ao redor do mundo, não se ilimitado mais a sua territorialidade e, dessa forma, tornando as plataformas locais públicos para debates e manifestações de seus usuários, não mais espaços para apenas compartilhar fotos e vídeos.

As mídias sociais se tornaram uma espécie de governadores desses espaços sociais compartilhados, uma vez que detêm o poder de criar e modificar regras e políticas internas, principalmente no que diz respeito à liberdade de expressão e aos direitos de privacidade. Esse processo acaba gerando obstáculos para que os usuários consigam, de forma igualitária, articularem-se para a proteção e garantia de seus direitos, uma vez que a plataforma elabora seus regramentos priorizando suas necessidades particulares [e muitas vezes voltadas à obtenção e aumento dos lucros da empresa]. E mesmo existindo um conjunto de regramentos de governança corporativa, os quais buscam assegurar direitos humanos fundamentais, esses direitos possuem um valor demasiadamente grande para que sejam estabelecidos de forma privada em um contrato de “termo de uso”. (Suzor, 2018, p. 4)

Ressalta-se que quando as empresas de tecnologia tomam decisões que afetam seus usuários, há poucas vias de reparação para as pessoas que sentem que foram tratadas injustamente. A lei federal dos EUA, por exemplo, fornece às empresas de tecnologia imunidade para suas decisões de moderar suas redes e as absolve da responsabilidade pelo que seus usuários dizem *online*. Seu poder de controlar seus usuários é protegido pela Primeira Emenda, mas a Primeira Emenda não protege os usuários das decisões das empresas de tecnologia. A Primeira Emenda apenas impede que o governo dos EUA interfira no discurso. As *bigtechs* norte-americanas são entidades privadas e são livres para decidir se prestam ou não serviços a uma

⁹ “[...] não há consenso sobre até que ponto os utilizadores necessitam de proteção contra as decisões de governação das próprias plataformas”. (Suzor, 2018, p. 4, tradução livre)

determinada pessoa ou grupo e não são obrigadas, sob a Primeira Emenda, a respeitar os direitos de liberdade de expressão dos usuários. (Suzor, 2019, p. 7-8)

É evidente que seria muito mais fácil que os grandes comandantes das mídias sociais ignorassem por completo qualquer interesse ou anseio de seus usuários ou não quisesse se adequar aos regramentos externos para evitar qualquer tipo de responsabilização em razão das extrapolações de limites que podem ocorrer em suas plataformas.

Mesmo que a Primeira Emenda estadunidense garanta sua isenção diante de qualquer ato indevido realizado por seus usuários, em outras localidades do mundo isso pode não ocorrer. Inclusive, vislumbra-se que esse entendimento de que as redes sociais não devem responder pelos danos causados dentro de suas comunidades está começando a mudar.

Justo por isso que a garantia e proteção de direitos fundamentais é um debate que não pode ser exercido e aplicado de forma unilateral pelas mídias sociais, haja vista que se inclui nessa controvérsia questões relativas à censura, discriminação algorítmica, abuso e assédio consumado nas plataformas, entre outras temáticas. Entretanto, a movimentação da sociedade em torno da regulamentação desses novos direitos tem produzido efeitos para pressionar os proprietários das *big techs*, refletindo em constantes mudanças nos “termos de usos” das plataformas. Porém, ainda é necessário assegurar os direitos dos usuários de uma forma mais imparcial, fora da plataforma, para que essa luta não seja barrada por questões contratuais corporativas. (Suzor, 2018, p. 4)

Lessig (2006), quando trata da governança na internet, trouxe uma ideia de que o “código” é a lei do mundo digital, isto é, os engenheiros que criam toda a estrutura da internet e suas plataformas, possuem o poder de controlar a vida digital [quem pode entrar, quem deve sair, o que pode ser dito, o que não pode ser dito, o que podemos assistir, o que não será nos mostrado...].

O código de *software* pode até ser certo, mas ele não é neutro. Todas as escolhas feitas pelas incorporações da internet possuem implicações políticas, segundo o ponto de vista de Lessig (2006), uma vez que as regras construídas pelas *big techs* não são menos políticas que as leis promulgadas democraticamente pelas casas legislativas de cada Estado-nação.

As concepções de Suzor (2019, p. 91) permite complementar que as empresas de tecnologia governam nossas vidas de maneira mais direta, através

dessas regras e processos internos, influenciando o comportamento de seus usuários nos espaços compartilhados. A ferramenta utilizada para governar o ciberespaço não são leis promulgadas democraticamente por atores legislativos, mas sim regras estabelecidas por meio de termos contratuais de serviços, entre outras normativas internas dessas empresas. Possuindo, assim, recursos técnicos para restringir o comportamento da comunidade, os quais são vantajosos para si ao serem suprimidos ou reforçar valores de acordo com suas necessidades. (Suzor, 2019, p. 91)

É nessa direção que há uma urgência, de determinados grupos, para que se institua um constitucionalismo digital, o qual percorre “*to rethink how the exercise of power ought to be limited (made legitimate) in the digital age*”¹⁰, com o intuito de proteger valores e direitos fundamentais através de uma governança descentralizada das *big techs*, retirando das plataformas o poder de execução de regras e medidas inerentes ao poder público (Suzor, 2018, p. 4; Gregorio, 2021, p. 57).

Ademais, assim como o processo que vem ocorrendo na União Europeia, a importância de efetivação dos novos direitos através do constitucionalismo digital é preencher a lacuna deixada pela inércia legislativa e desobrigar o poder judiciário de atuar de maneira ativista em matérias que necessitam de uma positividade jurídica. A tendência, a partir do modelo europeu, é que levar o constitucionalismo digital para uma fase transnacional, ou seja, “*whose aim is to protect fundamental rights and democratic values in the algorithmic society*”¹¹ (Gregorio, 2021, p. 66-70).

Enquanto a instauração do constitucionalismo digital para uma regulamentação ultraterritorial mais eficiente, igualitária e justa aplicada às mídias sociais ainda é um projeto abstrato, tem-se a figura dos intermediários para suprir, de certa forma, essa lacuna legislativa.

Suzor (2019, p. 90) trouxe em sua obra a ilustração dos intermediários, que são empresas que executam uma importante função na administração dos conteúdos que podem ser compartilhados e visualizados na internet, pois são esses *hots* de conteúdo, mecanismos de pesquisa e provedores de infraestrutura que definem as regras a serem aplicadas no meio virtual. Suas escolhas de regulamentação possuem um efeito concreto sobre os atos dos usuários nas mídias sociais, uma vez que

¹⁰ “[...] repensar como o exercício do poder deve ser limitado (legitimado) na era digital” (Suzor, 2018, p. 4, tradução livre)

¹¹ “[...] cujo objetivo é proteger os direitos fundamentais e os valores democráticos na sociedade algorítmica” (Gregorio, 2021, p. 70, tradução livre)

possuem a autoridade de delimitar o que pode ser pesquisado, visto e com quem os usuários podem se comunicar. Sua governabilidade advém da “*architecture of their networks*” e não de uma legislação, sendo, dessa forma, atores-chave do comando da vida das comunidades virtuais.

Todavia, em razão da centralidade dessa função, os intermediários podem estar suscetíveis a “*influence of many others, both governments and private entities, that want to see intermediaries regulate our behavior for their own ends, this is the tangled web of governance of the digital age*”¹². (Suzor, 2019, p. 90)

Sendo assim, por mais que esses novos atores no cenário da regulamentação e governança da internet e das mídias sociais estejam desassociados das grandes empresas de tecnologias, elas podem ser persuadidas por aqueles que detém um maior poder governamental ou econômico, para que passem a servir seus próprios interesses comerciais, os distanciando de sua suposta e esperada neutralidade, a qual se dá por meio de:

[...] social norms they encourage, through direct statements about what the platform is for, and by intervening to enforce rules in particular ways that deter behavior they want to suppress and reinforce values they want to encourage.¹³ (Suzor, 2019, p. 92)

O problema, porém, está na dificuldade em saber quais são os valores e comportamentos que esses indivíduos, por detrás de *hots* de conteúdo, mecanismos de pesquisa e provedores de infraestrutura, vão escolher para reforçar e limitar os comportamentos dentro do ciberespaço, haja vista que suas escolhas não estão pautadas em nenhum ordenamento jurídico, mas sim em suas próprias crenças e convicções.

Entretanto, pode-se completar que a partir da visão de Fachin (2021, p.8), que as normativas envolvendo o ciberespaço deveriam ser:

[...] adotadas para regular os conflitos interpessoais seriam produzidas pela convergência de vontades dos Estados nacionais. Nessa perspectiva, haveria uma instância internacional, com a participação dos Estados nacionais, para aprovar declarações, tratados, pactos e convenções em

¹² “[...] influência de muitos outros, tanto governos como entidades privadas, que querem ver os intermediários regularem o nosso comportamento para os seus próprios fins, esta é a teia emaranhada de governação da era digital”. (Suzor, 2019, p. 90, tradução livre)

¹³ “[...] normas sociais que eles incentivam, por meio de declarações diretas sobre a finalidade da plataforma e intervindo para fazer cumprir regras de maneiras específicas que dissuadam o comportamento que desejam suprimir e reforçam os valores que desejam encorajar.” (Suzor, 2019, p. 92, tradução livre)

termos idênticos ao que já é tradicional no âmbito desses organismos internacionais, como são os casos da ONU e da OEA. Esse modo de regulação evitaria alguns problemas, tais como o conflito de lei no espaço e, ao mesmo tempo, não haveria regiões do mundo sem regulação dos conflitos. Contudo, para que um modelo assim funcionasse efetivamente, seria necessário criar instituições com jurisdição universal. O direito do ciberespaço seria um produzido exclusivamente para reger as relações do mundo virtual, desvinculado do direito do mundo real. Seria um direito elaborado pelas instâncias competentes dos organismos internacionais, como a ONU e a OEA, que aprovariam documentos específicos sobre os temas do espaço virtual. (Fachin, 2021, p. 8)

Conseqüentemente, volta-se ao ponto de que urge uma normatização acerca dos limites das liberdades coletivas e individuais a ser empregada no meio virtual. Se essa escolha se dará por meio de um constitucionalismo digital ou se será eleito algum outro tipo de regulamentação na seara do direito internacional? Não há como sabermos.

Há quem seja contra a ideia do constitucionalismo digital, pois não acredita que o direito constitucional seja o instrumento correto para solucionar essa falha jurídica. Pode ser que nessa caminhada se descubra uma nova premissa legal, um novo nicho no âmbito jurídico ou até uma adequação a alguma área já existente.

Entretanto, o que se tem certeza é que as evoluções das tecnologias da informação e comunicação revolucionaram a sociedade mundial, mas com esses avanços tecnológicos, novas violações de direitos surgiram e, conseqüentemente, a busca por uma proteção desses novos direitos se torna impreterível para salvaguardar aqueles que tenham sido vítimas de transgressão. Nessa pauta foram categorizados os direitos humanos de 5ª geração, bem como a tentativa de se solidificar o constitucionalismo digital, principalmente com um caráter ultraterritorial, que proteja a sociedade mundial dos danos e violações causadas pelos impactos globalizados que as redes sociais e o acesso à internet causaram ao mundo.

3 A LIBERDADE DE EXPRESSÃO, SEUS LIMITES E OS PADRÕES DE COMUNIDADE DO META

Desde seu nascimento o ser humano busca se expressar de alguma forma. Balbucia, chora, gesticula, aprende a falar sílabas, palavras soltas, formula frases, aprender a escrever, quem não pode se expressar pela voz usa a linguagem de sinais. A música, o canto, a escrita, uma pintura, uma charge, uma fotografia. Expressar-se é um ato intrínseco do ser humano, é uma ação que caracteriza a singularidade da humanidade.

Ser livre para expressar o que se sente ou pensa é um direito garantido nas mais diversas nações ao redor do mundo, mas em contrapartida, suas restrições e limites dependem do entendimento e da legislação aplicados no território em que se está inserido. Entretanto, antes de se adentrar na questão dos limites e restrições à liberdade de expressão, é necessário compreender o que significa a liberdade de expressão.

Para Meyer-Pflug (2009, p. 66-70), a liberdade de expressão pode ser entendida como uma “exteriorização do pensamento, ideias, opinião, convicções, bem como de sensações e sentimentos em suas mais variadas formas, quais sejam, as atividades intelectuais, artísticas, científicas e de comunicação”, abrangendo diversas outras liberdades, as quais são protegidas pelo ordenamento jurídico de cada país, tais como a liberdade religiosa, de informação, imprensa, podendo se dar “por meio da escrita, ou uso de imagens e não necessariamente pessoalmente”.

Todavia, há quem entenda, também, que a liberdade de expressão pode ser dividida em duas dimensões, sendo a primeira associada à individualidade do ser humano, enquanto ser social, de se manifestar sem que haja uma censura prévia, independentemente da relevância que seu conteúdo tenha para o “interesse público ou para qualquer outra pessoa”, podendo ser expressada pela fala, arte ou escrita, inclusive através de “vestimentas, adornos ou tatuagens”. Enquanto a segunda dimensão diz respeito à liberdade de expressão em se constituir como um “ pilar da democracia, no sentido de sua indispensabilidade para a participação de indivíduo enquanto cidadão que pode criticar, opinar e atuar politicamente na comunidade na qual está inserido”. (Occiuzzi, 2017, p.18)

Assim, o chamado *free speech* está compreendido não apenas pelo direito do cidadão se expressar livremente, pelas mais diversas formas de manifestação, mas

pelo aspecto da democracia, haja vista que o direito ao voto e ao da participação na vida política, também são formas de expressão dentro de determinadas sociedades, o que define o cerne da democracia em si.

E, assim, pode se dizer que a liberdade de expressão resulta em um “conjunto de direitos relacionados às liberdades de comunicação”, os quais fazem parte de direitos fundamentais que devem ser assegurados para a concretização da livre manifestação, tais como a liberdade de manifestar pensamento ou opinião, bem como o direito de emitir e receber informação. (Magalhães, 2008, p. 74; Tôrres, 2013, p. 62)

Tôrres complementou diversos outros exemplos de direitos que estão abrangidos pela proteção à liberdade de expressão, pelo menos a conhecimento da Constituição Brasileira:

Dessa maneira, é correto dizer que, conexos e intrínsecos à liberdade de expressão, encontram-se também outros direitos, como o direito de informar e de ser informado, o direito de resposta, o direito de réplica política, a liberdade de reunião, a liberdade religiosa etc. Por conseguinte, a concepção de liberdade de expressão deve ser a mais ampla possível, desde que resguardada a operacionalidade do direito. (Tôrres, 2013, p. 63)

Pode-se acrescentar, também, o entendimento de Howard (2019, p. 95), no qual elucidou que a liberdade de expressão possui um significado moral e, também, legal. No que diz respeito ao significado moral, pressupõe-se dizer que todos os indivíduos são livres para expressar suas opiniões, crenças e convicções, bem como se comunicarem uns com os outros. No entanto, esse requisito moral gera deveres negativos à sociedade em geral, pois o direito à liberdade de expressão não deve se sobressair ao direito fundamental de outrem. Porém, esses deveres não se direcionam apenas ao indivíduo, uma vez que traz o dever do Estado em proteger aqueles que forem atingidos pelos efeitos negativos ao exercício da liberdade de expressão. O segundo significado corresponde à positivação do direito à liberdade de expressão de determinada jurisdição.

O direito à liberdade de expressão, portanto, não somente garante ao cidadão o direito de livre manifestação, de receber a informação ou opinião, mas deixa a cargo do Estado o dever de proteger possíveis danos causados por essa liberdade através de cada ordenamento jurídico, os quais terão limitações mais flexíveis ou restritivas, de acordo com a localidade de sua vigência.

Justo por isso, este segundo capítulo tem como objeto buscar compreender quais são os padrões de liberdade de expressão no direito internacional à luz dos direitos humanos, trazendo ao contexto, também, os padrões aplicados às comunidades do Meta Inc, com foco na plataforma *Facebook* e, por fim, contextualizar o direito à liberdade de expressão no direito norte-americano, para que se possa, enfim, realizar a análise proposta no objetivo geral desta pesquisa.

3.1 Os padrões internacionais sobre liberdade de expressão e garantia dos direitos humanos na moderação de conteúdo nas mídias sociais

Diversas são as fontes do direito, as quais correspondem à origem da norma jurídica, seja ela material ou formal, e que serão concretizadas na sociedade através da legislação, da jurisprudência dos Tribunais e Cortes, da doutrina e, também, pelos costumes. O direito, ainda, é subdividido em diversas categorias, as quais em sua maioria se limitam à territorialidade de cada nação.

Mesmo havendo uma centralização no direito pátrio, pode-se dizer que o direito internacional teve início no século XIX, quando a comunidade global passou a aderir a tratados formais, deixando de lado a aplicação dos costumes como fonte do direito. Contudo, essa mudança paradigmática não gerou grande repercussão na configuração das diretrizes internacionais. (Zibetti; Pilau; Miranda, 2018, p. 114)

Zibetti, Pilau e Miranda (2018, p. 114) expuseram que à medida que ocorria o “alargamento territorial das fronteiras” fez com que o direito, até então concentrado nas relações de determinados grupos europeus, fossem se expandindo tanto às nações de toda a América quanto ao extremo oriente. Segundo os autores, ainda não está claro como se originou esse processo de abrangência temática globalizada, mas podem ser indicados alguns contextos históricos, tais como:

[...] a revolução industrial e a revolução francesa como marcos dessa evolução. No caso desta, seguindo uma dinâmica já visível no longínquo advento da Carta Magna Inglesa e na proclamação do *Bill of Rights*, consolidou-se a tendência que, no futuro, levaria à emergência do indivíduo como sujeito de direito internacional. No caso da primeira, o avanço tecnológico, que a partir de então não conheceria mais freio, reduziu as distâncias intensificando as relações internacionais, mas também os problemas delas advindos. (Zibetti; Pilau; Miranda, 2018, p. 115)

Assim, a partir do momento em que passaram a surgir os primeiros obstáculos em negociações ou relações transnacionais no século XIX, a comunidade

global compreendeu que não conseguiria resolver essas questões sozinha, da mesma forma que os ordenamentos jurídicos pátrios não serviriam para auxiliar na solução de todos os conflitos que poderiam vir a surgir. (Zibetti; Pilau; Miranda, 2018, p. 116)

Diante desse novo prisma, ao longo dos anos, foram surgindo novas diretrizes e organizações internacionais, inicialmente relacionados ao sistema financeiro e comercial:

Num contexto de unificação dos padrões monetários, a aceleração da atividade financeira e a correlativa intensificação do risco sistêmico fizeram com que as autoridades monetárias nacionais desenvolvessem mecanismos próprios para ajustar os desequilíbrios das balanças de pagamentos, o que não evitou que ações coordenadas se fizessem necessárias em momentos de turbulência: algumas crises chegaram a exigir a adoção de medidas mais contundentes, como a concessão de empréstimos diretos de um banco central a outro. Em setores mais técnicos, observou-se, já naquela época, algumas iniciativas pioneiras de institucionalização da cooperação internacional, como a International Telegraph Union, criada em 1865, a Universal Postal Union, em 1875, a International Metereological Union, em 1878. (Zibetti; Pilau; Miranda, 2018, p. 116)

Entretanto, segundo Zibetti, Pilau e Miranda (2018, p. 116), foi após a Segunda Guerra Mundial que “a institucionalização da cooperação internacional mostrou-se enfim essencial”, principalmente com a criação da Organização das Nações Unidas (ONU), que passou a internalizar os conflitos transnacionais, como organização e instrumento para dirimir as tensões pós-guerra. E nessa nova ótica, as mais diversas organizações internacionais passaram a surgir, não restando outra alternativa, senão o reconhecimento de uma personalidade jurídica internacional:

[...] sem dúvida um marco conceitual na história do direito internacional porque se adicionava formalmente à sociedade internacional uma entidade distinta dos Estados, caracterizando-se com isso num elemento disruptor de um sistema pautado pela soberania. (ZIBETTI; PILAU; MIRANDA, 2018, p. 116)

Em se tratando de direitos humanos, podemos dizer que, com a preocupação para estabelecer um ordenamento jurídico de abrangência universal, a solidificação da Declaração Universal dos Direitos Humanos, pela ONU, em 1948, foi considerada um novo marco histórico, na qual se preocupou em positivar normativas que assegurassem à população mundial a garantia de direitos humanos fundamentais, tendo em vista as diversas violações ocorridas no período da Segunda Guerra Mundial pelo regime nazista.

Neste sentido, a Declaração Universal de Direitos Humanos (ONU, 1948), buscou garantir, no texto do artigo 19, a liberdade de expressão como um direito

humano e fundamental, o qual assegura a liberdade de transmitir, procurar e receber opiniões, ideias e manifestações por qualquer meio de comunicação, sem que houvesse qualquer interferência.

E ao buscar garantir a liberdade de expressão, não há no texto da DUDH a imposição de nenhuma norma limitadora desse direito, admitindo intrinsecamente a possibilidade de ocorrerem violações diante de uma normativa tão abrangente. Dessa forma, em 1966, a ONU promulgou o Pacto Internacional dos Direitos Civis e Políticos, no qual, em seus artigos 19 e 20¹⁴ (ONU, 1966), foram estabelecidos os limites da liberdade de expressão, incluindo, a proibição da apologia ao ódio. Em 1969, a Organização dos Estados Americanos, também, estabeleceu orientações sobre os limites da livre manifestação individual e coletiva através de seu artigo 13¹⁵ (OEA, 1969).

Segundo o documento publicado pela ONU, referente à Sessão 74 da Assembleia Geral, a liberdade de expressão é um direito fundamental, mas suas restrições devem ter um caráter excepcional, ou seja, os limites à livre manifestação não podem prejudicar o direito em si. Ademais, o ônus de apresentar a justificativa de

¹⁴ Artigo 19 §1. Ninguém poderá ser molestado por suas opiniões. §2. Toda pessoa terá o direito à liberdade de expressão; esses direitos incluirá a liberdade de procurar, receber e difundir informações e ideias de qualquer natureza, independentemente de considerações de fronteiras, verbalmente ou por escrito, de forma impressa ou artística, ou por qualquer meio de sua escolha. §3. O exercício de direito previsto no § 2 do presente artigo implicará deveres e responsabilidades especiais. Consequentemente, poderá estar sujeito a certas restrições, que devem, entretanto, ser expressamente previstas em lei e que se façam necessárias para: 1. assegurar o respeito dos direitos e da reputação das demais pessoas; 2. proteger a segurança nacional, a ordem, a saúde ou a moral públicas.

Artigo 20 §1. Será proibida por lei qualquer propaganda em favor da guerra. §2. Será proibida por lei qualquer apologia ao ódio nacional, racial ou religioso, que constitua incitamento à discriminação, à hostilidade ou à violência. (ONU, 1966)

¹⁵ Artigo 13 - Liberdade de pensamento e de expressão 1. Toda pessoa tem o direito à liberdade de pensamento e de expressão. Esse direito inclui a liberdade de procurar, receber e difundir informações e ideias de qualquer natureza, sem considerações de fronteiras, verbalmente ou por escrito, ou em forma impressa ou artística, ou por qualquer meio de sua escolha. 2. O exercício do direito previsto no inciso precedente não pode estar sujeito à censura prévia, mas a responsabilidades ulteriores, que devem ser expressamente previstas em lei e que se façam necessárias para assegurar: a) o respeito dos direitos e da reputação das demais pessoas; b) a proteção da segurança nacional, da ordem pública, ou da saúde ou da moral públicas. 3. Não se pode restringir o direito de expressão por vias e meios indiretos, tais como o abuso de controles oficiais ou particulares de papel de imprensa, de frequências radioelétricas ou de equipamentos e aparelhos usados na difusão de informação, nem por quaisquer outros meios destinados a obstar a comunicação e a circulação de ideias e opiniões. 4. A lei pode submeter os espetáculos públicos a censura prévia, com o objetivo exclusivo de regular o acesso a eles, para proteção moral da infância e da adolescência, sem prejuízo do disposto no inciso 2. 5. A lei deve proibir toda propaganda a favor da guerra, bem como toda apologia ao ódio nacional, racial ou religioso que constitua incitamento à discriminação, à hostilidade, ao crime ou à violência. (ONU, 1969)

limitação ao referido direito deve recair sobre a autoridade que impôs a restrição, não devendo repassar a responsabilidade ao interlocutor do discurso dizer se tem direito ou não de se manifestar daquela maneira. (ONU, 2019, p. 5)

Dessa maneira, ao impor uma limitação à liberdade de expressão, o órgão julgador deve atender 3 critérios, sendo eles: legalidade, legitimidade, necessidade e proporcionalidade. O primeiro critério compreende que para haver a limitação do direito, esses limites devem estar previstos na legislação de forma “precisa, pública e transparente”, não podendo ser abrangida de forma generalizada, bem como o interlocutor deve ser notificado sobre as razões pelas quais sua liberdade está sendo restringida. O segundo requisito obriga a autoridade que restringir o discurso ou manifestação em justificar o ato nos fundamentos do §3º, do artigo 19, do PIDCP. E, por último, a “necessidade e proporcionalidade” vai impor à autoridade que limitar o direito à liberdade de expressão demonstrar ser necessária a restrição desses direitos, “*to protect a legitimate interest and to be the least restrictive means to achieve the purported aim*”¹⁶ (ONU, 2019, p. 5).

Nessa mesma direção, a Organização dos Estados Americanos publicou em 2010 um marco jurídico, pelo qual estabelece os limites da liberdade de expressão a partir de uma perspectiva dos Padrões Interamericanos. No texto da obra, a Comissão alega que o sistema interamericano de direitos humanos é o ordenamento internacional que melhor reúne garantias para liberdade de expressão no mundo, tendo sido solidificado no artigo 13, da CADH. As garantias previstas no referido artigo trazem uma importante proteção ao direito humano à liberdade de expressão, não podendo serem utilizadas restrições que se não sejam aplicáveis na conjuntura interamericana, isto é, as restrições mais severas impostas à liberdade de expressão que não estão previstas no CADH, não poderão ser aplicadas nas decisões da Corte Interamericana de Direitos Humanos. No entanto, por se utilizar do princípio *pro homine* em seus julgamentos, a norma que prevalecerá será aquela que for mais favorável à pessoa humana (OEA, 2010, p. 1-2).

Isso se dá, justamente, pelo entendimento de que a liberdade de expressão é um dos direitos que mais identifica o ser humano como ele é: “*la virtud única y preciosa de pensar al mundo desde nuestra propia perspectiva y de comunicarnos*

¹⁶ “[...] para proteger um interesse legítimo e ser o meio menos restritivo para atingir o objetivo pretendido”. (ONU, 2019, p. 5, tradução livre)

*com los outros*¹⁷ e a partir do direito à livre manifestação é que outros direitos fundamentais, individuais ou coletivos, serão identificados. Outro panorama, pelo qual a CIDH respalda seus julgados é que a liberdade de expressão representa uma gama de direitos que são estruturais para o bom funcionamento da democracia e uma premissa insubstituível para a preservação desse estado democrático de direito, tais como: *“la formación de una opinión pública informada y consciente de sus derechos, el control ciudadano sobre la gestión pública y la exigencia de responsabilidad de los funcionarios estatales, no sería posible si este derecho no fuera garantizado”*¹⁸. (OEA, 2010, p. 2-3)

A liberdade de expressão se encontra no cerne da proteção dos direitos humanos, visto que é a partir dela que outros direitos fundamentais podem ser exercidos pela sociedade. Segundo entendimento da CIDH, a liberdade de expressão deve ser garantida pelos Estados ou pelas normas e diretrizes internacionais, caso contrário, a falta de proteção à livre manifestação incorrerá em violação de diversos outros direitos fundamentais (OEA, 2010, p. 4).

Todavia, é importante ressaltar que o exercício do direito à liberdade de expressão pressupõe aos interlocutores a imposição de deveres e responsabilidades sobre suas manifestações, sejam elas escritas ou verbalizadas. Sendo assim, aquele que exterioriza seus pensamentos e opiniões possui o dever de não violar o direito – ou exercício do direito – de outrem através de seu discurso. Da mesma forma que suas responsabilidades deverão ser analisadas dependendo da situação em que o direito à liberdade de expressão for exercido, bem como do modo empregado para externalizar e divulgar suas manifestações. A análise desses deveres e responsabilidades pela CIDH se dá por meio do cumprimento de 3 requisitos relativos à: legalidade, legitimidade, necessidade e proporcionalidade. (OEA, 2010, p. 6-24).

Basicamente o mesmo método aplicado para limitar a liberdade de expressão na CIDH é o mesmo empregado pela ONU nos julgados de sua Corte, os quais já foram explicados.

Todavia, com o avanço tecnológico e a utilização das redes sociais como uma ferramenta amplificadora do direito à liberdade de expressão, pode-se constatar

¹⁷ “[...] A virtude única e preciosa de pensar no mundo desde nossa própria perspectiva e de nos comunicarmos uns com os outros”. (OEA, 2010, p. 3, tradução livre)

¹⁸ “[...] a formação de uma opinião pública informada e consciente de seus direitos, o controle cidadão sobre a gestão pública e a exigência de responsabilidade dos funcionários estatais, não seria possível se este direito não fosse garantido”. (OEA, 2010, p. 3, tradução livre)

que, conforme explanado no subtópico “Da necessidade de novas configurações para a proteção dos direitos humanos na Era Digital”, ainda não há uma norma ultraterritorial que regule a moderação de conteúdo na internet, isto é, não há uma legislação internacional mais abrangente para a temática no ciberespaço.

Na atualidade, a liberdade de expressão e a moderação de conteúdo nas mídias sociais, em regra, vem sendo tratada conforme o ordenamento jurídico de cada Estado-nação, fazendo com que a legislação já existente se adapte às mudanças que a Era Digital trouxe ao cotidiano jurídico ou crie novas normas para melhor atender as necessidades dos conflitos no ciberespaço e na vida *online*, o que se pode ter como alguns exemplos: o Marco Civil da Internet, no Brasil; a seção 230 do Communications Decency Act (CDA), nos Estados Unidos; a *Netzwerkdurchsetzungsgesetz* (NetzDG), na Alemanha, podendo ser traduzida como Lei de Fiscalização da Rede; a Lei Avia, na França; o *Online Harms White Paper*, no Reino Unido; e a Carta Portuguesa de Direitos Humanos na Era Digital, em Portugal.

Entretanto, é importante ressaltar que em tratando da questão da regulamentação das mídias sociais, Archegas *et al.* (2021, p. 8) explica que desde seu surgimento, as plataformas digitais haviam se consolidado em uma “esfera de autorregulação” e que essa autorregulação poderia ser traduzida “na estipulação de termos de uso e padrões da comunidade que guiam a atividade de moderação de conteúdo em serviços como buscadores e redes sociais”.

Ou seja, antes de se desencadear um movimento global pela regulamentação da moderação de conteúdo, a Seção 230, do CDA, foi aprovada nos Estados Unidos, fazendo com que a partir da metade da década de 90, as mídias sociais e plataformas digitais em geral passassem a “atuar dentro de um paradigma de autorregulação”. Essa autorregulação consiste no estabelecimento de normas internas, as quais podem ser identificadas como “Padrões da Comunidade” ou “Termos de Uso”, e que irão modelar o desenvolvimento dos algoritmos para a moderação de conteúdo dentro da plataforma. (Estarque; Archegas, 2020, p. 19)

Pode-se observar que o preceito da autorregulação atribui aos provedores intermediários uma autonomia ilimitada para o estabelecimento de diretrizes para a moderação de conteúdo de seus usuários, sendo assim, neste modelo o Estado não tem o poder de intervir na estruturação das regulamentações de funcionamento das plataformas, ficando a cargo das empresas a estipulação da definição de padrões de uso por seus usuários.

Foi somente após 2010 que esse “paradigma da autorregulação passou a ser questionado”, a sociedade e poder público passaram a forçar uma mudança deste modelo, de modo que Estarque e Archegas (2020, p.19) explicaram que:

Diversos países passaram a pressionar as empresas donas das redes sociais mais influentes, exigindo que elas fossem mais eficientes no combate à desinformação online. Isto é, elas deveriam atuar de maneira mais intensa e moderar não apenas posts obscenos ou ofensivos, mas também conteúdos mais controversos como publicações deliberadamente falsas ou com o intuito de enganar outros usuários (Estarque; Archegas, 2020, p. 19)

Os autores ainda elucidam que na “era do bem-estar público” passou-se a se preocupar mais com os impactos “dos danos gerados pelo discurso de usuários nas redes sociais” e de que forma as mídias sociais estavam lidando com o problema e buscando soluções eficientes para o controle de conteúdo dentro do modelo de autorregulação. Citou-se como exemplo o escândalo da *Cambridge Analytica* e o *Facebook* [caso este já abordado no capítulo anterior], o qual intensificou e evidenciou o problema da desinformação e da manipulação algorítmica nas redes sociais. Vale ressaltar que os problemas e prejuízos não se limitam a apenas esses dois fenômenos, a Era Digital e o crescimento do ciberespaço acabaram por intensificar diversos outros fenômenos como: “discurso de ódio, bullying, disseminação de pornografia não consensual ou *revenge porn*, violação de direitos autorais, violação da esfera de proteção de crianças e adolescentes, discriminação, dentre outras”. (Archegas *et al.*, 2021, p. 8)

Assim, mesmo após o Meta Inc. buscar uma solução efetiva com a criação e implementação do *Oversight Board*, comitê este que será melhor explicado no último capítulo desta pesquisa, as implicações profundamente alarmantes dos danos causados pela disseminação de conteúdos nocivos e violadores de direitos humanos continuaram a atormentar a sociedade em geral, uma vez que deixar a supervisão da moderação de conteúdo apenas em poder das *bigtechs* pode trazer uma sensação de insegurança jurídica.

À vista disso, retornando ao âmbito do direito internacional, em 2020 a Comissão Europeia apresentou uma proposta inovadora de regulamentação das mídias sociais e de seus conteúdos compartilhados no ciberespaço: o *Digital Services Act* (DSA). O referido ato legislativo, o qual está em vigor desde novembro de 2022, mas terá efetivamente aplicada a totalidade de suas obrigações a partir de 2024. A normativa consiste em “criar um mercado único para serviços digitais dentro dos

limites da União Europeia”, tendo em seu escopo a garantia de proteção dos direitos fundamentais e a contribuição para a segurança online, ou seja, uma tentativa de proteção aos usuários das mais diversas plataformas digitais por meio da imposição de deveres aos provedores de serviços. (Archegas *et al.*, 2021, p. 6)

De acordo com Archegas *et al.* (2021, p. 6-7), o DSA conta com a aplicação de princípios de governança corporativa, tais como *accountability* e a transparência, ferramentas essas que facilitam a fiscalização das plataformas digitais pelos órgãos públicos da União Europeia. Ainda segundo os autores, esse monitoramento é feito por meio de um conselho geral, chamado de *European Board for Digital Services*, que dará o suporte necessário para a fiscalização exercida pelos coordenadores de cada Estado-membro, chamados de *Digital Services Coordinators*.

Dessa forma, para que ocorra a aplicação efetiva dos deveres de transparência e responsabilização, de acordo com DSA, as plataformas intermediárias deverão disponibilizar anualmente um relatório de transparência (*transparency reports*), no qual prestarão informações relativos aos dados acerca da aplicação de medidas moderadoras de conteúdos em seus provedores, “including the measures taken as a result of the application and enforcement of their terms and conditions”¹⁹. (Archegas *et al.*, 2021, p. 6-7; Comissão Europeia, 2020, p. 13)

É importante destacar que o DSA não buscar substituir a autorregulação das *big techs* e impor em seu lugar uma regulamentação de regência estatal, o que a legislação europeia traz é a imposição de “alguns parâmetros básicos de transparência que, por sua vez, irão garantir de uma forma mais eficiente a proteção do interesse público dentro da esfera digital, hoje com um espaço mais amplo de controle por parte das empresas privadas”. (Archegas *et al.*, 2021, p. 7)

O objetivo da regulamentação é, portanto, uma forma de buscar uma maior segurança aos usuários do ciberespaço e garantir a aplicação e a proteção de direitos fundamentais entre todos os atores desse processo, incluindo, também, os provedores intermediários numa esteira de responsabilidade da fiscalização do que acontece dentro de suas plataformas:

[...] a ideia é compatibilizar o modelo de autorregulação com a proteção do interesse público com base na estruturação de um modelo de co-regulação. [...] É justamente nessa terceira fase da governança de plataformas que se inserem as propostas hoje debatidas na União Europeia e no Reino Unido. A

¹⁹ “[...] incluindo as medidas tomadas como resultado da aplicação e execução dos seus termos e condições”. Comissão Europeia, 2020, p. 13, tradução livre)

partir da atribuição de competências a uma organização independente, esses novos arranjos regulatórios buscam garantir a prevalência do interesse público na Internet sem que isso signifique o fim da autorregulação de plataformas. (Archegas, 2021, p. 9)

Inequívoco dizer, portanto, que a liberdade de expressão vem sendo resguardada como direito fundamental no direito internacional desde o fim da Segunda Guerra Mundial, mas com a eclosão da Sociedade da Informação e Comunicação, da Era Digital e da transformação espaço-temporal que o ciberespaço sofreu com a Quarta Revolução Industrial, a proteção do *free speech* e a moderação de conteúdo do que é postado na internet carece, ainda, de uma regulamentação internacional mais abrangente.

Muito embora cada Estado-nação esteja, ainda, se adaptando a esses impactos e buscando soluções legislativas para esses problemas, a grande diversidade de decisões as quais recaem ou recairão em uma mesma plataforma são incontáveis. Ademais, os danos causados por postagens ofensivas nas redes sociais estão cada vez mais amplificados e ultrapassando limites territoriais, o que exige que a comunidade internacional busque soluções mais efetivas.

Claro que a ONU e a OEA através de suas cortes buscam uma aplicação justa e equilibrada para a limitação da liberdade de expressão e não se pode mitigar o importante papel de seus julgamentos de competência internacional, seus parâmetros de legalidade, legitimidade, necessidade e proporcionalidade também influenciaram nos critérios de julgamento de moderação de conteúdo no *Oversight Board*, da mesma forma que esses parâmetros estão dispostos, também, na DSA.

A DSA, de fato, é uma regulamentação inovadora, mas possui uma aplicação restrita aos países que fazem parte do bloco econômico da União Europeia. Mas como será a recepção deste novo ordenamento jurídico? Como as plataformas irão se portar ao longo do tempo com as exigências do poder público em relação aos impactos causados no ambiente virtual?

Pode ser que a DSA, como um novo padrão internacional de direitos humanos, aplicados ao ciberespaço e suas mídias sociais, realmente tenha uma efetividade no combate aos danos causados pela disseminação em massa de conteúdos violadores de direitos fundamentais. Pode ser, também, que a regulamentação seja fonte inspiradora para outras fontes normativas do direito internacional que abranjam outras nações. Ou, ainda, pode ser que seja constatado que a regulamentação precisa de alterações para seu melhor funcionamento, pois o

estabelecido inicialmente não é suficiente para os fins pelos quais foi criada.

Todavia, enquanto ainda não se tenha resposta, por meio de dados concretos, acerca da efetividade da DSA na fiscalização da moderação de conteúdo nas redes sociais, a comunidade jurídica internacional continuará buscando soluções para os conflitos existentes no mundo digital os quais ultrapassam os limites da territorialidade de cada Estado-nação.

3.2 A autorregulação dos provedores intermediários, Termos de Uso e os padrões da comunidade do *Meta Inc.*

Conforme desenvolvido até o presente momento, pode-se observar que desde a promulgação do ato normativo da Seção 230 da CDA, nos Estados Unidos, deu-se às plataformas digitais o direito à autorregulação, isto é, é dado aos provedores intermediários a prerrogativa de criarem suas próprias normas, sem que o Estado interfira na elaboração e execução de suas regras de funcionamento.

Entretanto, a aprovação da CDA, em 1996, não foi tão bem recebida pelos defensores do direito à liberdade de expressão, tal qual John Perry Barlow, que acreditava em uma regulamentação orgânica do ciberespaço, ou seja, a internet deveria ser um espaço independente do Estado e despreendido do mundo real, no qual haveria de ser muito mais democrático, pois a maneira de governar e implementação de regras se daria dentro desse espaço com a participação de seus usuários. (Barlow, 1996)

E, para além, segundo Barlow, a internet deveria ser um espaço livre de limitações à liberdade de expressão:

Cyberspace consists of transactions, relationships, and thought itself, arrayed like a standing wave in the web of our communications. Ours is a world that is both everywhere and nowhere, but it is not where bodies live. We are creating a world that all may enter without privilege or prejudice accorded by race, economic power, military force, or station of birth. We are creating a world where anyone, anywhere may express his or her beliefs, no matter how singular, without fear of being coerced into silence or conformity. Your legal concepts of property, expression, identity, movement, and context do not apply to us. They are all based on matter, and there is no matter here.²⁰
(Barlow, 1996)

²⁰ O ciberespaço consiste em transações, relacionamentos e no próprio pensamento, dispostos como uma onda estacionária na rede de nossas comunicações. O nosso mundo é um mundo que está em todo lugar e em lugar nenhum, mas não é onde os corpos vivem. Estamos a criar um mundo onde todos podem entrar sem privilégios ou preconceitos concedidos por raça, poder econômico, força militar ou posição de nascimento. Estamos a criar um mundo onde qualquer pessoa, em qualquer

Talvez para depois da metade da década de 90, com o advento e a solidificação do ciberespaço, o ideal de Barlow parecia fazer sentido. A busca por um espaço mais democrático, no qual poderia ser regido por aqueles que se utilizavam desse ambiente. Todos em comum acordo e coabitando de forma pacífica. Um espaço de respeito mútuo e, principalmente, um espaço que acima de tudo respeita a liberdade de cada indivíduo se manifestar da maneira que acredita estar correto.

Suzor (2019, p. 88-89) trouxe sua crítica ao ideal de Barlow, pois muito embora os libertários do ciberespaço buscassem ratificar o princípio da “*neutrality of the 'net*”, com uma possível imposição de censura e vigilância do poder público sobre a internet, essa ideologia de que a “tirania da maioria forçava a minoria a se submeterem a regras que não aceitavam” e de que a comunidade virtual seria muito bem capaz de criar suas próprias regulamentações de uma maneira justa e democrática, nada mais era que uma visão utópica desse movimento.

Em outras palavras, Barlow não poderia prever que o espaço virtual fosse ser, futuramente, dominado por pequenos grupos, os quais detém o poder de controle de quase a totalidade dos meios virtuais. Hoje, a saída de um membro do ciberespaço, por não concordar com suas regras, não gera nenhum impacto efetivo para melhora de seus ordenamentos internos, os quais são muito mais fáceis de se adaptarem de forma mais efetiva às necessidades dos usuários, do que leis construídas democraticamente e fora do ambiente virtual. (Suzor, 2019, p. 89)

O que Suzor quis demonstrar com sua crítica é que Barlow não sabia o que estava por vir com os avanços tecnológicos que passaram a surgir, principalmente, a partir do início do 2º milênio. A evolução do ciberespaço para uma vida *onlife* na atualidade, a concentração do poder nas mãos de uma minoria que possui o controle das *bigtechs*, isso fez com que os impactos causados dentro do espaço virtual ultrapassassem ao mundo real. Logo, ao causar repercussão fora do espaço virtual, não é de se surpreender que o problema passe a fazer parte do interesse público.

E isso Suzor (2019, p. 3) identificou em sua obra, na qual esclarece que apesar de que as regulamentações das mídias sociais, por meio dos “termos de uso”,

lugar, pode expressar as suas crenças, por mais singulares que sejam, sem medo de ser coagida ao silêncio ou à conformidade. Os seus conceitos legais de propriedade, expressão, identidade, movimento e contexto não se aplicam a nós. Todos eles são baseados na matéria e não há matéria aqui. (Barlow, 1996, tradução livre)

tragam a proibição de abusos praticados por seus usuários, “essas plataformas não fazem o suficiente para impor suas regras”. Mas, com o crescimento da pressão do poder público e das organizações da sociedade civil, as intermediárias se viram forçadas a tomarem atitudes, as quais:

*[...] responding to increasing pressure by more clearly articulating their standards of acceptable behavior and banning users and groups that spread hatred and abuse. These rules are not yet uniformly enforced, but they are becoming enforced more regularly.*²¹ (Suzor, 2019, p. 3)

Ressalta-se que antes de iniciar a movimentação do Estado e da sociedade pela regulamentação das mídias sociais, os provedores de serviços não se envolviam em debates públicos pautados no conteúdo que seus usuários postavam em suas plataformas, e isso se dava porque essas empresas entendiam pela neutralidade diante de seus usuários. Mesmo assim, cada fornecedor de serviços online possui a prerrogativa de decidir o que é e o que não é permitido falar, discutir e debater dentro de suas comunidades, o qual se dá por meio de um contrato costumeiramente chamado de “Termos de Uso”, mas que pode ter outras denominações também. (Suzor, 2019, p.4)

Esses contratos consistem nos regramentos do que pode e o que não pode ser feito dentro da cada plataforma, incluindo cláusulas que permitem a rescisão do contrato a qualquer momento por parte do provedor. Todavia, quando essas empresas de tecnologia passam a tomar decisões que afetam os usuários de suas redes, ao entendimento da autorregulação *“there are few avenues of redress for people who feel that they have been treated unfairly”*²², nas palavras de Suzor (2019, p. 7), o autor trouxe como exemplo a aplicação da legislação norte-americana nos provedores com sede nos Estados Unidos:

US federal law provides technology companies immunity for their decisions to moderate their networks, and absolves them of liability for what their users say online. Their power to control their users is protected by the First Amendment, but the First Amendment does not protect users from the decisions of technology companies. The First Amendment only prevents the US government from interfering with speech; US tech companies are private

²¹ “[...] respondendo à pressão crescente articulando mais claramente os seus padrões de comportamento aceitável e banindo utilizadores e grupos que espalham ódio e abuso. Estas regras ainda não são aplicadas de maneira uniforme, mas estão a ser aplicadas com maior regularidade.” (Suzor, 2019, p. 3, tradução livre)

²² “[...] há poucas vias de reparação para as pessoas que sentem que foram tratadas injustamente”. (Suzor, 2019, p. 7, tradução livre)

*entities and are free to decide whether or not they provide services to a particular person or group. US tech companies are not obliged, under the First Amendment, to respect free speech rights of users.*²³ (Suzor, 2019, p 7-8)

A partir desse entendimento, cada provedor poderia aplicar em seus “termos de uso” proibições que vão de encontro aos seus interesses em consonância com a lei local, se esta lhe for favorável. Suzor (2019, p. 6) argumentou que a diferença da legislação de cada país poderia influenciar os provedores a escolherem uma jurisdição que lhes fosse mais benéfica para sua própria proteção, isto é, sem que houvesse riscos de terem suas plataformas banidas ou suspensas.

Lessig (1999), em sua obra que trata da governança na internet, trouxe uma ideia de que o “código” é a lei do mundo digital. Isto é, os engenheiros que criam toda a estrutura da internet e suas plataformas, possuem o poder de controlar a vida digital (quem pode entrar, quem deve sair, o que pode ser dito, o que não pode ser dito, o que podemos assistir, o que não será nos mostrado, e etc.). Entretanto, ele faz a ressalva de que o código de *software* pode até ser certo, mas ele não é neutro, uma vez que todas as escolhas feitas pelas plataformas possuem implicações políticas, ou seja, as regras construídas pelas *big techs* não são menos políticas que as leis promulgadas democraticamente pelas casas legislativas de cada país. (Lessig, 1999)

E isso trouxe um alerta na última década, pois ficou cada vez mais visível que, muito embora os pequenos grupos que controlam os maiores meios digitais não sejam as entidades mais poderosas do mundo, estes são os atores-chave do controle da vida de seus usuários, uma vez que a centralização do controle das mídias sociais os tornam suscetíveis a influência de tantos outros influentes que possuem o desejo de regular o comportamento de sua comunidade para alcançar seus próprios objetivos. (Suzor, 2019, p. 90)

Suzor (2019, p. 91), ainda, descreveu que essa ferramenta utilizada para governar o ciberespaço, para o controle de comportamento dos usuários, não são leis promulgadas democraticamente por atores legislativos, mas sim regras estabelecidas por meio de termos contratuais de serviços, entre outras normativas internas dessas

²³ “A lei federal dos EUA concede imunidade às empresas de tecnologia pelas suas decisões de moderação das suas redes e as isenta de responsabilidade pelo que os seus usuários dizem online. O seu poder de controlar os seus usuários é protegido pela Primeira Emenda, mas a Primeira Emenda não protege os usuários das decisões das empresas tecnológicas. A Primeira Emenda apenas impede que o governo dos EUA interfira no discurso; as empresas tecnológicas dos EUA são entidades privadas e são livres de decidir se prestam ou não serviços a uma determinada pessoa ou grupo. As empresas tecnológicas dos EUA não são obrigadas, ao abrigo da Primeira Emenda, a respeitar os direitos de liberdade de expressão dos utilizadores.” (Suzor, 2019, p. 7-8, tradução livre)

empresas. Possuindo, assim, recursos técnicos para restringir o comportamento da comunidade, os quais são vantajosos para si ao serem suprimidos ou reforçar valores de acordo com suas necessidades.

Inequívoco dizer que ao longo dos anos, os olhares do interesse público, diante das implicações que a manipulação algorítmica e a disseminação de discursos violadores de direitos humanos, se voltaram para não somente a moderação de conteúdo, mas também para a forma em que é realizada a autorregulação das plataformas digitais.

Com isso, pode-se dar como exemplo de aprimoramento das regulamentações internas o *Meta Inc.*, a qual sofreu grandes modificações desde sua criação. Assim, cada vez mais a plataforma acrescenta regras para o bom funcionamento de sua comunidade, bem como para a moderação do conteúdo postado em suas redes sociais por seus usuários.

De acordo com o próprio *site* do *Meta Inc.*, a rede social preza pela livre manifestação de seus mais de 2 bilhões de usuários, mas não tolera qualquer tipo de abuso proveniente da liberdade de expressão e, portanto, atribui a si o papel de eliminar conteúdos que vão contra os padrões da comunidade, baseados em orientações de especialistas das áreas de tecnologia, segurança pública e direitos humanos (Facebook).

Ainda, segundo o *site*, o compromisso da plataforma é dar voz aos seus usuários, mas quando estes extrapolam os limites impostos pela comunidade, o direito à liberdade de expressão é limitado por meio da avaliação de quatro valores:

- a. Autenticidade: garantia de um conteúdo autêntico, não podendo utilizar a plataforma para falsificar sua própria identidade ou o que se está fazendo;
- b. Segurança: garantia de um lugar seguro para se expressar, sendo vedada postagens contendo ameaças, intimidações, exclusão ou silenciamento de pessoas;
- c. Privacidade: garantia de proteção e privacidade das informações pessoais de seus usuários;
- d. Dignidade: garantia de um ambiente pautado no respeito à dignidade alheia, sendo proibido o assédio e a difamação de terceiros. (Facebook)

Tendo sido avaliado o conteúdo e identificada uma postagem supostamente violadora, é necessário identificar em quais das categorias que está sendo infringidos os “Padrões da Comunidade”, que são divididas da seguinte forma:

1. Violência e comportamento criminoso: violência e incitação, organizações perigosas e indivíduos perigosos, coordenação de danos e divulgação de crime, produtos controlados, fraude e dolo;
2. Segurança: automutilação e suicídio, exploração sexual, abuso ou nudez infantil; exploração sexual de adultos, bullying e assédio, exploração humana, violação de privacidade e direitos de privacidade imagem;
3. Conteúdo questionável: discurso de ódio, violência e conteúdo explícito, nudez adulta e atividade sexuais, abordagem sexual;
4. Integridade e autenticidade: integridade da conta e identidade autêntica, spam, segurança cibernética, comportamento não autêntico, notícias falsas, mídia manipuladora, perfil memorial;
5. Respeito à propriedade intelectual;
6. Solicitações e decisões relativas a conteúdo: solicitações de usuários, proteção adicional de menores de idade. (Facebook)

A partir das categorias apresentadas nos “Padrões de Comunidade”, pode ser destacada a subcategoria “Organizações e indivíduos perigosos”, a qual é a categoria em que o foi identificada a infração de Donald Trump, ao ter suas contas suspensas tanto no *Facebook* quanto no *Instagram*.

A referida categoria é uma política que busca evitar e acabar com os danos no mundo real, proibindo qualquer manifestação diretamente de organizações ou indivíduos perigosos ou manifestações de apoio ou incitação à violência. Sendo dividida em 3 níveis para classificação do conteúdo:

- a. O **Nível 1** concentra-se em entidades que se engajam em sérios danos no meio físico, incluindo a organização ou a defesa da violência contra civis, a desumanização repetida ou a defesa de danos contra pessoas com base em características protegidas ou o envolvimento em operações criminosas sistemáticas. As entidades de Nível 1 incluem organizações terroristas, criminosas e de ódio. Removemos elogios, apoio substancial e representação de entidades de Nível 1, bem como de seus líderes, fundadores ou membros notáveis. O Nível 1 inclui organizações de ódio; organizações criminosas, incluindo aquelas designadas pelo governo dos Estados Unidos como Comandos Especiais de Tráfico de Drogas (SDNTKs, pelas iniciais em inglês); e organizações terroristas, incluindo entidades e indivíduos designados pelo governo dos Estados Unidos como Organizações Terroristas Estrangeiras (FTOs, pelas iniciais em inglês) ou Terroristas Globais Especialmente Designados (SDGTs, pelas iniciais em inglês). Removemos elogios, apoio substancial e representação de entidades de Nível 1, bem como de seus líderes, fundadores ou membros notáveis.
- b. O **Nível 2** concentra-se em entidades que se envolvem em violência contra agentes estatais ou militares, mas que geralmente não visam civis, o que chamamos de "Agentes violentos não estatais". Removemos todo apoio substancial e representação dessas entidades, bem como de seus líderes e membros notáveis. Removemos qualquer elogio às atividades violentas desses grupos.
- c. O **Nível 3** concentra-se em entidades que podem se envolver repetidamente em violações de discurso de ódio ou das Políticas de Organizações Perigosas dentro ou fora da plataforma ou demonstrar forte intenção de se envolver em violência no meio físico futuramente, mas que não necessariamente se envolveram em violência até o momento ou defenderam a violência contra outros com base nas características protegidas deles. Isso inclui movimentos sociais militarizados, redes de conspiração de

violência e indivíduos e grupos banidos por promoverem o ódio. As entidades de Nível 3 podem não ter presença ou coordenação nas nossas plataformas. (Facebook)

O *Meta Inc.*, também, não permite que seus usuários promovam conteúdos que contenham elogios, apoio substancial ou de representação de autores de ataques terroristas, eventos de ódio, atos criminosos com várias vítimas ou tentativas destes atos, bem como de assassinatos em série ou crimes e conteúdo de ódio, tais quais o nazismo e supremacia branca.

Cumprir destacar que o conglomerado de plataformas compreende que compartilhar conteúdo da categoria “Indivíduos e Organizações Perigosas” com intuito de denúncia, condenação do ato ou abertura para discussão de forma neutra, não violam os padrões da comunidade, reforçando que as políticas da rede “são criadas para dar espaço para esses tipos de discussão e, ao mesmo tempo, limitar os riscos de possíveis danos no meio físico”.

E diante do exposto, Negócio (2022, p. 4) trouxe uma consideração importante da temática debatida neste subtópico, na qual consiste que a autorregulação das plataformas digitais pode servir tanto para a prevenção ou garantia de “uma boa imagem da estrutura privada” como, de certa forma, busca afugentar uma possível intervenção do poder público ao se adequar às exigências legais impostas posta pelos Governos. Isto dá às *bigtechs* uma segurança para continuar atuando com autonomia a partir do momento em que passa a cooperar com os interesses públicos.

Pode-se observar, portanto, que ao passo que o poder público e a sociedade civil passaram a se interessar pela forma como as mídias sociais estavam controlando seus conteúdos, a partir do irrefutável impacto que o conteúdo postado em suas plataformas estava gerando dentro e fora do ciberespaço, as empresas tiveram que remodelar suas regras internas, tendo que se adequar às mais diversas legislações, para que abrangesse às exigências não somente do mercado, mas às demandas estatais também e com isso, esquivava-se de uma intervenção legal mais severa que retirasse sua autonomia de autorregulação.

Inclusive a criação do *Oversight Board*, pelo *Meta Inc.*, faz parte das medidas tomadas pela empresa para a contenção de danos causados por postagens nocivas e para a adequação de governanças aplicadas aos direitos humanos no debate da moderação de conteúdo.

Contudo, verifica-se que mesmo que tenha ocorrido essa movimentação “espontânea” da plataforma, ainda há muita controvérsia sobre interferência e autonomia do Tribunal do *Facebook* perante à rede sociais a qual está vinculada, mas isso não impede que se mapeie e analise os prós e contras de sua atribuição, no qual o objeto da presente pesquisa propõe. Mas, claro, a partir da limitação do tema apresentado ao caso da suspensão de contas do ex-presidente Donald Trump.

3.3 Os limites da liberdade de expressão nos Estados Unidos

Inicialmente, antes de se adentrar no cerne deste subcapítulo, é interessante fazer uma ressalva do motivo pelo qual foi levantada a proposta de compreender a liberdade de expressão norte-americana.

Diferentemente do que se compreende a liberdade de expressão no Brasil ou, até mesmo em parte, com o entendimento do *free speech* trazido pelo direito internacional, conforme elucidado no subcapítulo anterior, o conceito norte-americano de uma liberdade irrestrita demonstra sua contribuição para os atos que ocorreram até o fatídico “6 de janeiro” e que culminou na suspensão de contas do ex-presidente Donald Trump.

Nos Estados Unidos, a Primeira Emenda²⁴ trouxe em seu texto a liberdade de expressão como alicerce do estado democrático de direito norte-americano, proibindo de forma expressa qualquer cerceamento do direito a emitir opiniões seja através do povo, por meio de manifestações, ou da imprensa, não podendo o Estado proibir o cidadão norte-americano de se manifestar ou criticar, independentemente do conteúdo ser ofensivo ou não, uma vez que isso o afastaria do debate público, sendo incompatível – e, portanto, inconstitucional – com a garantia à liberdade de expressão (Stone, 1987, p. 55; 1994, p.81).

A idealização de um direito à liberdade de expressão absoluto vem da compreensão de que esta seria uma condição “*sine qua non for political legitimacy in a democracy*”²⁵ e segundo Waldron (2012, p. 173-175), nessa lógica, o *free speech*

²⁴ “O congresso não deverá fazer qualquer lei a respeito de um estabelecimento de religião, ou proibir o seu livre exercício; ou restringindo a liberdade de expressão, ou da imprensa; ou o direito das pessoas de se reunirem pacificamente, e de fazerem pedidos ao governo para que sejam feitas reparações de queixas.” (Estados Unidos, 1791, tradução livre)

²⁵ “[...] *sine qua non* para a legitimidade política em uma democracia”. (Waldron, 2012, p. 173, tradução livre)

só contribuiria para o sentimento de democracia se ela não possuísse restrições, caso contrário, seria deslegitimada. O autor ainda expôs que os defensores da ampla liberdade, compreendem que, como sociedade, a maioria não possui o direito de impor restrições a uma minoria, não sendo justo estabelecer restrições, sem que tenham a oportunidade de expressar suas opiniões, atitudes e, inclusive, preconceitos e ideologias.

Além disso, uma das inquietações dos opositores à regulamentação se manifesta pelo medo de serem julgados ou até processados por aquilo que decidem externalizar informalmente com pessoas próximas:

*Opponents of hate speech regulation sometimes say that these laws are targeted at what people can say in bars or at the dinner table, and occasionally they cite examples of people being prosecuted for what they thought they were saying just among friends.*²⁶ (Waldron, 2012, p. 240)

Claro que esse entendimento advém muito em razão da cultura de determinada sociedade, tal qual a norte-americana que se está abordando neste capítulo, que no ímpeto de sua liberdade de expressão irrestrita, o cidadão sabe que não há problema em se proferir um discurso de ódio, porque mesmo sendo um discurso repudioso, ele está amparado constitucionalmente. Waldron (2012, p. 176), inclusive, levantou essa questão, também, que o sistema legislativo e político de cada sociedade é determinado por seu ambiente moral e cultural, fazendo com seus hábitos e costumes locais influenciem na criação de leis e ampara os discursos políticos e inseridos nesse contexto estão não somente as opiniões, mas também preconceitos, discriminações e intolerância.

E assim, quando se trata do *free speech* norte-americano, este traz a problemática da proteção irrestrita, fazendo com que os interlocutores do discurso nocivos aos direitos humanos se sintam encorajados a emitirem opiniões preconceituosas, racistas, misóginas, xenofóbicas, homofóbicas e, inclusive, incitar movimentos antidemocráticos sem que seus atos sejam punidos mais severamente perante a justiça.

Minow (2000, p. 1260-1262) já trazia aos seus estudos os argumentos dos defensores da Primeira Emenda utilizavam contra a regulamentação do discurso de

²⁶ “Os opositores à regulamentação do discurso de ódio dizem por vezes que estas leis visam o que as pessoas podem dizer em bares ou à mesa de jantar e, ocasionalmente, citam exemplos de pessoas que foram processadas por aquilo que pensavam estar a dizer apenas entre amigos”. (Waldron, 2012, p. 240, tradução livre)

ódio, para eles, quando se abre uma brecha legislativa para limitar a liberdade de expressão, faz com que cada vez mais exceções possam ser aplicadas à lei e com isso, prejudicar a própria liberdade individual do cidadão. Em contrapartida, os defensores da regulamentação arguem que a falta dela resulta em um imenso dano individual e, também, à sociedade como um todo:

Thus, too often, the advocates of the First Amendment ignore or try to minimize the ways in which slurs and bias-based comments both produce psychological damage for individuals and perpetuate the dehumanization of members of particular groups (which in turn can invite further degradation and violence). Freedom of speech that undermines equality neglects the way that equality is itself a central principle of the First Amendment. [...] Speech absolutists ought at least to worry about the net reduction, and content distortion, of speech's marketplace that hate speech can produce. Arguably, even worse than this kind of "silencing" is what happens to those who nonetheless persist and speak. [...] That protection for freedom of expression and thought is or should be absolute. Yet, it is only honest to acknowledge the harms first, and then decide what to do. Of course, other harms have justified restrictions on speech. Consider how far from absolute protection is our existing law. It permits restrictions for civil actions for defamation and slander; it permits restrictions against fighting words and words that amount to a clear and present danger to national security or physical safety.²⁷ (Minow, 2000, p. 1260-1262)

Minow (2000, p. 1262) ainda admitiu que a restrição à liberdade de expressão instituída de forma específica ao discurso de ódio pode ser conciliada às liberdades e garantias individuais dentro de um Estado Democrático de Direito, visto que o *hate speech* tem a capacidade de afetar outras liberdades individuais. A autora ainda rebate que diversos países enfrentam o tema através de restrições específicas e que buscar conhecer o modo como se dá essa restrição seria muito didático aos juízes norte-americanos, se esses estivessem dispostos a aprender com outras vertentes constitucionais.

²⁷ “Assim, com demasiada frequência, os defensores da Primeira Emenda ignoram ou tentam minimizar as formas como os insultos e os comentários tendenciosos produzem danos psicológicos para os indivíduos e perpetuam a desumanização de membros de grupos específicos (o que, por sua vez, pode provocar maior degradação e violência). A liberdade de expressão que mina a igualdade negligencia a forma como a igualdade é em si um princípio central da Primeira Emenda. [...] Os absolutistas do discurso deveriam pelo menos se preocupar com a redução líquida e a distorção do conteúdo do mercado do discurso que o discurso de ódio pode produzir. Indiscutivelmente, ainda pior do que este tipo de “silenciamento” é o que acontece com aqueles que, no entanto, persistem e falam. [...] Essa proteção à liberdade de expressão e pensamento é ou deveria ser absoluta. No entanto, é honesto reconhecer primeiro os danos e depois decidir o que fazer. É claro que outros danos justificaram restrições à expressão. Consideremos quão longe da proteção absoluta está a nossa lei existente. Permite restrições para ações civis por difamação e calúnia; permite restrições contra palavras de combate e palavras que representam um perigo claro e presente para a segurança nacional ou física.” (Minow, 2000, p. 1260-1262)

E esse é um ponto bastante relevante quando se está defronte com a jurisprudência estadunidense, a qual nitidamente defende a liberdade irrestrita e não parece conseguir – ou busca conseguir – equilibrar os conflitos constitucionais entre a liberdade de expressão e a dignidade da pessoa humana.

Contextualizando a partir das décadas de 1950 e 1960, período em que movimentos sociais afloraram pelos Estados Unidos, muitos deles com o propósito de denunciar as chagas do racismo e das leis segregacionistas no país, conhecido como o movimento dos direitos civis. Os debates acerca dos limites à liberdade de expressão passaram a pautar ações e julgamentos nos tribunais e na Suprema Corte, podendo ser separados alguns casos emblemáticos, como: *Brandenburg vs. Ohio* (1969), *National Socialist Party vs. Skokie* (1977) e *R.A.V vs. City of St. Paul* (1992) (Ribeiro; Costa, 2016, p. 151).

Os limites da liberdade de expressão tiveram início pelos Tribunais Estadunidenses com a teoria *clear and present danger*, que surgiu a partir do caso *Schenck vs. United States*. A teoria consistia que o discurso que tivesse capacidade de criar um “perigo claro e presente”, daria ao Congresso o direito de tomar a atitude de restringi-lo por meio de processo legislativo. Ademais, o direito constitucional à liberdade de expressão – assim como outros direitos fundamentais – deveriam ser interpretados de formas diferentes quando uma nação está em tempos de paz ou quando está em tempos de guerra, podendo neste último caso, serem limitados severamente. (Lynd, 1975, p. 153-154)

Segundo Lynd (1975, p. 153-154) a referida doutrina foi adotada pelos tribunais norte-americanos com a entrada dos Estados Unidos na 1ª Guerra Mundial em resposta à oposição socialista e, justo por isso, o julgamento de cada ato dependeria do contexto em que ocorresse e esse entendimento era muito aplicado quando se tratava da defesa do Estado contra atos de “falas dos comunistas”. A partir disso, o governo norte-americano poderia penalizar qualquer cidadão que defendesse alguma ação ilegal, mesmo que este ato ainda viesse ocorrer em um futuro não definido, mas não poderia puni-lo se esta defesa estivesse vinculada um ato abstrato, o qual não virá a ocorrer.

Dessa forma, a teoria do perigo claro e presente fica evidente para a compreensão da aplicação de limites à liberdade de expressão nesse primeiro momento, ou seja, só haveria o cerceamento do *free speech* se a conduta defendida significasse em uma ameaça real naquele momento, mesmo que a resultado concreto

e certo se desse em um outro momento, a defesa desse ato significava um perigo no presente e, portanto, era suficiente para que o Estado tomasse alguma atitude para evitar o resultado danoso.

Entretanto esse entendimento foi reinterpretado na década de 50, através do caso *Dennis v. United State*, no qual consistia que a aplicabilidade da teoria não precisaria mais que o perigo estivesse presente, mas sim deveria ser questionada a probabilidade ou não da gravidade do ato, para que assim pudesse justificar a restrição da liberdade de expressão para obstar o perigo que o discurso pudesse causar, isto é, mesmo que em consequência do discurso não fosse concreta, a remota possibilidade de futuramente existir um grave perigo já ensejaria na limitação. (Lynd, 1975, p. 154-155)

Muitas foram as interpretações e reinterpretações da *clear and present danger theory* nos tribunais estadunidenses e na Corte Superior, mas a teoria só teve sua modificação significativa após o julgamento do caso *Brandenburg v. Ohio*, em 1969.

O caso foi julgado na Suprema Corte norte-americana, a qual decidiu que Clarence Brandenburg, líder da Ku Klux Klan, poderia manifestar seu discurso discriminatório, em uma entrevista televisionada, contra a população negra e, inclusive, ameaçar o Congresso caso continuasse a suprimir os direitos dos brancos em um comício da KKK²⁸, sendo estabelecido que o Estado não poderia se eximir de atuar contra o discursos que extrapolassem os limites à liberdade de expressão, desde que essa defesa tenha o objetivo de restringir a incitação ou produção de atos ilícitos iminentes, sendo assim, o teste de *Brandenburg* trouxe a teoria do “perigo iminente de um ato ilícito para que o discurso sofra uma limitação” (Brugger, 2007, p. 129).

Sem dúvidas, o teste de *Brandenburg* impactou a aplicação dos limites jurídicos da liberdade de expressão nos Estado Unidos, tendo refletido em casos como o *National Socialist Party vs. Skokie*, no qual se entendeu que a “possibilidade de realização de uma marcha de neonazistas, vestidos com uniforme da SS em um subúrbio de Chicago onde habitavam primordialmente sobreviventes do Holocausto”

²⁸ “Não somos uma organização vingativa, mas se o nosso Presidente, o nosso Congresso, o nosso Supremo Tribunal continuarem a suprimir a raça branca e caucasiana, é possível que seja necessária alguma vingança.” (Lynd, 1975, p. 152, tradução livre)

não eram razoáveis o suficiente para suprimir a liberdade de expressão daquele grupo, adotando assim, o princípio da neutralidade de conteúdo (Ribeiro; Costa, 2016, p. 150-151).

O princípio da neutralidade de conteúdo se tornou uma enorme aliado à essa proteção quase que ilimitada da liberdade de expressão nos Estados Unidos, fazendo com que interlocutores de discursos com profundo teor racista, preconceituoso, entre outras formas de violações aos direitos humanos, possam não ser sujeitos a uma penalização ou proibição pelo Estado (Alves; Ribeiro; Cardoso, 2012, p. 6), o que também ocorreu no julgado *R.A.V. vs. City of St. Paul*, quando um grupo de adolescentes queimaram cruzeiros no pátio de uma família negra, o Judiciário norte-americano decidiu que o Estado não poderia proibir um discurso de conotação e incitação racista ou sexista sem que proibisse, também, o discurso contra atos de racismo ou sexismo, pois dessa forma estaria “privilegiando os tolerantes em prol dos intolerantes” (Ribeiro; Costa, 2016, p. 151).

É possível, portanto, vislumbrar o excesso de proteção que o ordenamento jurídico norte-americano deu à liberdade de expressão após consolidação do teste de Brandenburg, fazendo com que os discursos mais profanos não possam ser proibidos ou censurados, para que haja um equilíbrio do debate de ideias e ideologias, salvaguardando um dos pilares do estado democrático de direito.

Mesmo que a Corte estadunidense tenha decidido por uma teoria mais irrestrita do *free speech*, entre os estudiosos, operadores do Direito e pesquisadores, o entendimento aplicado não possui uma anuência unânime, Minow, inclusive trouxe ao debate da temática da liberdade de expressão no artigo *Regulating hatred: whose speech, whose crimes, whose power? – an essay for Lenneth Karst*, os argumentos e desdobramentos daqueles que defendem um discurso irrestrito.

Segundo Minow (2000, p. 1271) para alguns a regulação do discurso de ódio e a possível restrição da liberdade de expressão pode causar um resultado negativo para a sociedade em geral, pois pode “desencadear novos e crescentes ressentimentos contra os supostos beneficiários das proteções”. De acordo com a professora, a partir desse posicionamento as pessoas que tivessem seus discursos restringidos pelo Estado poderiam se tornar “mártires e até garotos-propaganda”, podendo fortalecer uma cultura política que se aproveitaria de uma sensação de injustiça para se fortalecer politicamente, uma vez que “ *censorship and punishment of*

*ideas can render the ideas forbidden fruit, which are especially attractive to youth engaged in rebellion and others searching for symbols of disobedience*²⁹.

Bom, de certa forma, as consequências trazidas por Minow não são tão fora da realidade assim, como iremos ver na abordagem do caso de 6 de janeiro e a utilização desse discurso pelo então presidente Donald Trump para ascender politicamente e incitar seus fãs fervorosos contra a legitimidade do processo eleitoral norte-americano.

Ademais, Minow (2000, p. 1266-1267) complementou que a regulamentação ou limitação do *free speech* pode ser conduzido para uma possível aplicação seletiva da lei, na qual impactaria muito mais os grupos minoritários e fracos da sociedade, além de que poderiam resultar no crescimento de reações violentas em razão da aplicação díspar da legislação reguladora pelos grupos que sofrem com os maus-tratos dos discursos de ódio e de preconceito. Assim, os defensores do discurso irrestrito acreditam que criar limitações promoveria o aumento de discurso racistas, de diversos tipos de preconceito, dentre outros extremantes danosos, haja vista que a possibilidade de punição geraria revolta dos interlocutores e elevaria a violência:

*But hate speech proposals and hate crime laws do select for special regulation these kinds of group-based divisions rather than targeting qualities such as greed, political party affiliation, or countless other potential bases for disagreement and violence. Groups could as easily be defined by reference to these kinds of traits. Hence, laws that confine themselves to racial, sexual, religious, and ethnic groups are under inclusive. The rejoinder to this point is like the defense of antidiscrimination laws-some group traits are highlighted in order to remedy past and ongoing discrimination.*³⁰ (Minow, 2000, p. 1268)

Não somente Minow, mas o professor Edwin Baker trouxe ponderações semelhantes acerca da ineficácia à imposição de limites à liberdade de expressão, para o autor, para aqueles que são contra a regulamentação, esta funcionaria dentro de uma lógica de prevenção de danos e que a imposição da lei é muito simplista para

²⁹ “[...] a censura e a punição de ideias podem torná-las frutos proibidos, o que é especialmente atraente para jovens envolvidos em rebeliões e outros que procuram símbolos de desobediência” (Minow, 2000, p. 1271, tradução livre)

³⁰ Mas as propostas de discurso de ódio e as leis sobre crimes de ódio selecionam para regulamentação especial este tipo de divisões baseadas em grupos, em vez de visarem qualidades como a ganância, a filiação em partidos políticos ou inúmeras outras bases potenciais para desacordo e violência. Os grupos poderiam ser facilmente definidos por referência a esses tipos de características. Consequentemente, as leis que se limitam a grupos raciais, sexuais, religiosos e étnicos são subinclusivas. A réplica a este ponto é como a defesa de leis anti-discriminação – algumas características de grupo são realçadas a fim de remediar a discriminação passada e atual. (Minow, 2000, p. 1268, tradução livre)

modificar de forma eficaz o comportamento prejudicial daqueles que se utilizam da liberdade de expressão para causar danos a outrem. Da mesma forma em que a limitação que, em tese, serviria para prevenir os danos, poderia ser utilizada contra as próprias minorias, quem deveria ser defendido.

A esse respeito, pode-se pegar o exemplo que Minow (2000, p. 1265) utilizou em relação à expressão popular de “dois pesos e duas medidas”, na qual comparou o caso de *R.A.V. v. City of St. Paul* e *Wisconsin v. Mitchell*, para ela a aplicação do teste de Brandenburg foi aplicado de forma parcial nos dois casos, pois enquanto no primeiro caso havia um réu branco, este considerou inconstitucional a aplicação da pena de crime de ódio, caso esse que foi relatado acima.

Entretanto, no segundo caso, tendo um réu negro, o entendimento utilizado para a aplicação do teste foi utilizado de forma diversa, pois o Tribunal considerou que houve crime de ódio quando um jovem negro e seus amigos agrediram fisicamente um adolescente branco depois de discutirem sobre um filme da Ku Klux Klan que relatava uma violência contra uma criança negra orando³¹. (Minow, 2000, p. 1265)

No primeiro seu julgamento do caso, o júri entendeu que a lei punia o pensamento preconceituoso e não a conduta, nesse caso estava sendo punida a conduta de Mitchell em selecionar a vítima por sua raça/cor, mas a decisão foi revertida pela Suprema Corte de Wisconsin, na qual considerou que a aplicação de lei naquele caso estava punindo o pensamento preconceituoso e não a conduta, pois embora as regulamentações antidiscriminação ou de crimes de ódio reconheçam as implicações e danos que esses crimes podem causar, *“the right to free speech more important than the deterrence of such crimes”*³². (Trackston, 1993, p. 65)

O Estado recorreu da decisão e a Suprema Corte dos Estados Unidos reformou a decisão da Suprema Corte de Wisconsin, justificando a modificação da decisão a partir do entendimento de que *“the state is entitled to punish some forms of forbidden conduct more harshly than others if it determines such penalty enhancement*

³¹ O caso ocorreu em 1989, quando um grupo de jovens discutiu após assistir ao filme *Mississippi Burning*. A discussão de se por causa de uma cena em que um homem branco espanca uma criança negra que estava orando e, de acordo com o relato do ocorrido, após a sessão o grupo atacou um menino branco de 14 anos. No entanto Mitchell não participou da agressão, ao que foi alegado, o réu foi quem apontou para o menino. (Trackston, 1993, p. 65)

³² “O direito à liberdade de expressão é mais importante do que a dissuasão de tais crimes”. (Trackston, 1993, p. 65, tradução livre)

*is appropriate*³³, reconhecendo que crimes de ódio podem resultar em retaliação e agitação da comunidade local, além dos impactos negativos sobre as vítimas. (Trackston, 1993, p. 66-67)

Ainda, segundo Trackston (1993, p. 67) a decisão da Suprema Corte dos Estados Unidos admitiu uma posição de esperança para o combate dos crimes de ódio, mas a decisão foi duramente criticada, haja vista que os críticos das limitações à liberdade de expressão entendem que esse tipo de norma jurídica pode fazer uma associação perigosa à criminalização de qualquer tipo de discurso ou de pensamento.

Claramente a reforma do julgamento traz questões importantes para a aplicação ou não dos limites à liberdade de expressão nos Estados Unidos, tendo em vista que há uma mudança de interpretação entre os casos *R.A.V. v. City of St. Paul* e *Winsconsin v. Mitchell*, conforme foi mencionado o apontamento de Minow. Qual é a razão para a aplicação da restrição à liberdade de expressão nos dois casos? Estaria o tribunal aplicando um entendimento diferente para cada caso em virtude da raça dos réus? Ou foi porque os atos se deram de forma diferente? Em um caso a liberdade de expressão resultou em uma agressão física, mas no outro caso o discurso não ultrapassou as vias do *free speech*?

Estes questionamentos passam a perpetrar diante da análise desses casos, visto que a própria Suprema Corte norte-americana concorda que o Estado pode sim punir condutas proibidas mais severamente que outras. Mas qual é o marco para identificar o que pode ser punido e o que não pode ser punido?

Baker (1997, 989-990) argumentou que a fala muitas vezes pode causar danos a outrem, todavia, essa lesão se dá muito pela forma pela qual o ouvinte se sente atingido, seja por se sentir humilhado ou difamado ou até pela forma como irá reagir diante do discurso que lhe foi direcionado (se irá agir de forma criminosa ou imprudente:

This conception of autonomy means that when a person "offers" a view to another, the person only presents herself, exercises her autonomy. Any consequences involved in the listener's reaction or response, if the speaker's autonomy is protected and the listener's recognized, must be attributed, in the end, to the listener. The result is a right of the speaker to present her viewpoint even if its assimilation by the listener leads to or constitutes serious harm. But if liberty or autonomy is so important, why pick out certain liberties for special protection? Why protect a person's verbal expression of hatred rather than, for example, her freedom to express herself by throwing a stone at another

³³ "O Estado tem o direito de punir algumas formas de conduta proibida de forma mais severa do que outras, se determinar que tal aumento da pena é apropriado". (Trackston, 1993, p. 66, tradução livre)

person? Both are expressive and both can cause harm, at least as I have used the term. A hint at the difference has already been given in the discussion of "mental" elements, a point that will relate to a more careful look at the notion of liberty.³⁴ (Baker, 1997, p. 992)

O aludido por Baker corrobora com o entendimento aplicado pelo teste de Brandenburg, no qual consiste em trazer esse conflito do interlocutor do discurso e a lesão do ouvinte, colocando os dois lados em uma mesma balança para buscar identificar qual a conduta é mais potencialmente lesiva: o discurso lesivo do ofensor ou a restrição de seu discurso? No contexto norte-americano, a balança passa a pender sempre em favor da liberdade de expressão.

Baker (1997, p. 996), inclusive trouxe esse entendimento em seu texto, argumentando que por mais que um discurso possa causar efeitos negativos no ofendido, tais como *"a personal rejection that leaves a person emotionally unable to work or a truthful disclosure that causes a person to lose a job or an election"*³⁵, contudo, a identificação desses danos não poderia, *per se*, resultar na restrição da liberdade de expressão. Assim, somente quando o direito fundamental à liberdade de expressão não está em risco, é que se pode passar para a prevenção dos danos causados pelo discurso, ou seja, a restrição ou controle do discurso só pode ocorrer se o dano não advier do próprio exercício do direito à liberdade de expressão.

Não obstante toda a teoria e defensores da liberdade de expressão irrestrita, houve uma discussão bastante interessante entre o professor Ronald Dworkin e o juiz George Brunn, da Califórnia, publicada pela *The New York Review*, após Brunn criticar a decisão do caso *The New York Times Co. v. Sullivan*³⁶.

³⁴ "Esta concepção de autonomia significa que quando uma pessoa "oferece" um ponto e vista a outra, a pessoa apenas se apresenta, exerce a sua autonomia. Quaisquer conseqüências envolvidas na reação ou resposta do ouvinte, se a autonomia do falante for protegida e a do ouvinte reconhecida, devem ser atribuídas, no final, ao ouvinte. O resultado é o direito do falante de apresentar o seu ponto de vista, mesmo que a sua assimilação pelo ouvinte conduza ou constitua um dano grave. Mas se a liberdade ou a autonomia são tão importantes, por que escolher certas liberdades para proteção especial? Porquê proteger a expressão verbal de ódio de uma pessoa em vez de, por exemplo, a sua liberdade de se expressar atirando uma pedra a outra pessoa? Ambos são expressivos e podem causar danos, pelo menos como usei o termo. Uma sugestão da diferença já foi dada na discussão dos elementos "mentais", um ponto que se relacionará com um olhar mais cuidadoso sobre a noção de liberdade." (Baker, 1997, p. 992, tradução livre)

³⁵ "[...] uma rejeição pessoal que deixa uma pessoa emocionalmente incapaz de trabalhar ou uma revelação verdadeira que faz com que uma pessoa perca um emprego ou uma eleição" (Baker, 1997, p. 996)

³⁶ Mas se o Times ou uma rede ou eu acusarmos falsamente um funcionário de aceitar subornos, é inexplicável por que eles ou eu não deveríamos ser responsáveis pelos danos resultantes, especialmente quando a falsidade prejudica seriamente ou destrói a carreira de uma pessoa. (Dorkwin, 1992)

O texto consistia na crítica da reforma da decisão da Corte do Alabama, na qual acusou de reescreverem a lei de difamação daquele estado. Para ele a lei de difamação era extremamente necessária para “*discourage false statements of fact and to compensate the victims of such falsehoods*”³⁷, que a aplicação delas em nada afeta o direito à liberdade de expressão. Ademais, Brunn (1992) argumenta que embora a mídia em geral precise ter preservado seu direito à crítica e deva ter todos os direitos à liberdade de expressão ampliados para que não sofra com censura, não pode ela não ser responsabilizada por danos causados por falsas alegações e difamações veiculadas em suas publicações que afetem e causem danos a qualquer pessoa.

Dworkin (1992) por sua vez, respondeu às indagações de Brunn, e inicialmente destacou que a Suprema Corte estadunidense não fez nada a mais que seu papel de julgador, que mais cedo ou mais tarde a Corte teria que decidir sobre a constitucionalidade da lei de difamação do Alabama, mas que embora o Primeira Emenda determine que o Estado não possa restringir a liberdade de expressão ou de imprensa, a norma constitucional deve sempre ser interpretada antes de ser aplicada, dando como exemplo do caso *Brown v. Board of Education* e também o caso levantado pelo juiz em sua carta.

O professor prosseguiu na discussão avaliando que é melhor a interpretação abstrata da *First Admendment* pela Suprema Corte, do que ocorrer a revogação das interpretações anteriores se darem por meio de um processo falho de emendas à Constituição. Ademais, essa é a responsabilidade da Corte de “*para interpretar e proteger a Constituição*”, ou seja, sempre perceber um erro de interpretação, é papel dela corrigir quando provocada.

Logo, podemos perceber, que muito embora exista um estigma de que o tratamento à liberdade de expressão seja plenamente irrestrito nos Estados Unidos, muito por conta de que é um sentimento e interpretação majoritária na sociedade e cortes, há quem defenda que a liberdade de expressão possa ser interpretada conforme sua aplicação no caso concreto, a história jurídica envolvendo o *free speech* aparenta ter momentos tímidos de modificação na aplicação do teste de Brandenburg.

Entretanto, a interpretação da liberdade de expressão irrestrita continua sendo um posicionamento majoritário e é por essa interpretação que muitos dos

³⁷ “[...] desencorajar declarações falsas de fatos e compensar as vítimas de tais falsidades”. (Brunn, 1992, tradução livre)

interlocutores acreditam estar amparados, perpetuam discursos perigosos não somente a determinado grupo minoritário, mas também está causando riscos à democracia e seus processos eleitorais, tal como aconteceu nas últimas eleições presidenciais norte-americanas, que baseado nesse ideal do *free speech*, o ex-presidente Donald Trump se utilizou da prerrogativa de proteção dada pela Primeira Emenda para incitar dúvida e o ódio contra o processo eleitoral, no qual foi derrotado, e contra à democracia.

Este, portanto, é o tópico que será abordado no terceiro e último capítulo dessa pesquisa, em que será refeita uma recapitulação dos acontecimentos anteriores até a consumação dos atos de 6 de janeiro no Capitólio, para então analisar o papel do *Oversight Board* na manutenção da suspensão das contas de Trump nas plataformas digitais controladas pela *Meta Inc.*

4 ANÁLISE DO JULGAMENTO DAS CONTAS DE DONALD TRUMP PELO OVERSIGHT BOARD

4.1 O *Oversight Board*: criação, composição e funcionamento

Com o passar dos anos desde o surgimento das redes sociais; a evolução do ciberespaço, o crescimento do alcance da internet ao redor do mundo trouxe a esse espaço virtual uma maior preocupação e a necessidade de uma maior vigilância desses espaços, vistos os impactos que as publicações ou utilização das redes sociais têm causado para além de suas comunidades virtuais.

E não é de hoje que o interesse público pela moderação de conteúdo e a exigência de práticas mais responsáveis e transparentes têm forçado às *big techs* em buscar novos meios de dirimir esses conflitos e problemas em suas plataformas.

Segundo Klonick (2020, p. 2424-2425), por muito tempo as regras internas que determinavam a moderação de conteúdo nas redes sociais permaneceram desconhecidas de seus usuários. Cada vez mais as plataformas do *Meta Inc.* passaram a ser exigidos uma conduta mais transparente e que contassem com a participação de seus usuários não somente no processo criação de padrões, mas também na atualização e aplicação de regras para moderação de conteúdo.

Foi em abril de 2018 que Mark Zuckerberg começou a indicar publicamente que estava atento às reivindicações do poder público e da sociedade civil, indicando o início da elaboração do projeto para o *Oversight Board*, a partir de uma espécie de “Suprema Corte”, a qual seria independentemente da plataforma e teria a atribuição de realizar uma última análise e julgar sobre qual seria o discurso mais aceitável nas redes sociais. A ideia inicial era, também, que essas decisões refletissem em normas e valores mais globalizados. (Klonick, 2020, p. 2424-2425)

É importante ressaltar que muito embora a limitação à liberdade de expressão nas plataformas do *Meta Inc.* fosse bem menos complacente que o entendimento norte-americano da Primeira Emenda, a pressão sobre a plataforma se deu por meio das nações europeias, tais quais França e Alemanha, que insistiam na adaptação da imposição de limites à liberdade de expressão em cumprimento aos seus ordenamentos jurídicos nacionais. Ademais, ainda há indícios da colaboração do Sul Global para essas mudanças, visto que contam com a maioria dos usuários da

rede social e possuem uma vasta diversidade cultural e de entendimentos acerca dos limites à livre manifestação. (Klonick, 2020, p. 2437)

Entretanto, para Klonick (2020, p. 2438) mesmo havendo essa pressão da sociedade civil das mais diversas nações que utilizavam as plataformas do conglomerado de Zuckerberg, não foi suficiente para que seus diretores concretizassem uma maior publicidade e transparência a criação de regras e moderação de conteúdo:

The move to transparency operated as a tacit acknowledgment that Facebook users had a right to know the rules governing them and voice their reactions. But Facebook's reveal was not only insufficient to stem outrage; it in fact fueled further critique. The public's discovery that a small cadre of people headquartered in Silicon Valley were the sole creators and deciders on rules governing this vital global platform for online speech and that, although rules existed, their operation lacked core ideas of procedure and process added fuel to long-standing comparisons between Facebook and a feudal state, kingdom, or dictatorship.³⁸ (Klonick, 2020, p. 2438)

Não é surpresa, portanto, que o *Meta Inc.* só passou escutar com mais seriedade a necessidade da aplicação de transparência e do compartilhamento de como são feitos e aplicados os padrões da comunidade após “uma série de controvérsias geraram uma impressão negativa para a empresa e minaram a confiança do usuário”, os quais foram muito exemplificados por Klonick (2020, p. 2441):

O primeiro surgiu dos danos percebidos pelo Facebook à democracia e às eleições após a eleição presidencial dos EUA em 2016 e o “Brexit”, o referendo britânico para deixar a União Europeia. As alegações de que a interferência estrangeira por meio de notícias falsas e campanhas de desinformação afetou o resultado desses processos democráticos dominaram o ciclo de notícias e a opinião pública por meses. O segundo grande momento veio com o “vazamento” de dados do usuário por meio da consultoria terceirizada *Cambridge Analytica*. Em um comunicado à imprensa em 16 de março de 2018, o Facebook reconheceu que a empresa, administrada no Reino Unido, obteve acesso indevido e usou dados de oitenta e sete milhões de usuários do Facebook, com possíveis efeitos nas eleições americanas. Terceiro e mais recentemente, o Facebook foi acusado de aumentar a agitação civil e talvez até mesmo causar a morte de milhares de muçulmanos *Rohingya*, um grupo étnico perseguido em Mianmar, quando

³⁸ “A mudança para a transparência funcionou como um reconhecimento tácito de que os utilizadores do Facebook tinham o direito de conhecer as regras que os regem e de expressar as suas reações. Mas a revelação do Facebook não foi apenas insuficiente para conter a indignação; na verdade, alimentou novas críticas. A descoberta do público de que um pequeno grupo de pessoas sediadas no Vale do Silício foram os únicos criadores e decisores sobre as regras que governam esta plataforma global vital para o discurso online e que, embora existissem regras, a sua operação carecia de ideias centrais de procedimento e processo adicionou combustível a longo prazo. comparações permanentes entre o Facebook e um estado, reino ou ditadura feudal.” (Klonick, 2020, p. 2438, tradução livre)

grupos majoritários usaram o programa “Free Basics” do *Facebook* e aplicativos de mensagens (incluindo o *WhatsApp*) para mobilizar movimentos de limpeza étnica. A escala dessas questões e o clamor público resultante desencadearam uma nova pressão dos governos democráticos ocidentais. De 2017 a 2019, o Senado dos EUA realizou cinco audiências sobre o uso da tecnologia e seu impacto no discurso, eleições, terrorismo e privacidade. Isso incluiu o testemunho de Mark Zuckerberg perante o Senado após o escândalo da *Cambridge Analytica*, bem como o testemunho da COO Sheryl Sandberg após a eleição presidencial dos EUA em 2016 sobre a interferência estrangeira nas mídias sociais. Durante o mesmo período, a Câmara dos Deputados realizou quatro audiências, centralmente focadas nas práticas de filtragem de plataforma e na exploração do viés contra a moderação de conteúdo conservadora americana. (Klonick, 2020, p. 2441-2444, tradução livre)

E foi em 2018 que Zuckerberg falou pela primeira vez sobre o projeto do Comitê de Supervisão. Segundo o CEO do *Meta Inc.*, estava cada vez mais evidente que a plataforma não poderia deter a prerrogativa de “tomar decisões importantes sobre liberdade de expressão e segurança por conta própria”. Assim, os dirigentes da plataforma passaram a se reunir com diferentes grupos de pesquisadores e advogados para debater qual seria a melhor forma de concretizar esse projeto (Klonick, 2020, p. 2450)

O *Oversight Board* ou Comitê de Supervisão foi inaugurado em outubro de 2020 pode ser considerado um órgão julgador não-estatal, independente, que possui a atribuição de julgar postagens removidas, sejam elas de forma automática ou por meio de denúncias dos usuários, nas plataformas digitais do *Facebook* e do *Instagram*. Sua composição é disposta por 40 julgadores escolhidos como representantes de diversos territórios ao redor do mundo. (Oversight Board)

Os casos que chegam para julgamento no comitê devem ser feitos pelo usuário que teve seu pedido de reativação de postagem negado pela plataforma em que foi publicado o conteúdo removido. O usuário, portanto, terá quinze dias para apresentar sua apelação ao Comitê de Supervisão, o qual irá estudar o caso e verificar se ele se encaixa nos quatro requisitos de admissibilidade: difícil; significativo; de relevância global; e norteador de políticas futuras. Sendo assim, pode-se perceber que somente aqueles casos que se encaixem nas premissas de repercussão geral serão escolhidos para serem julgados. (Oversight Board)

Tendo sido selecionado o caso, este será distribuído para um subgrupo, o qual é chamado de “painel”, e será dado início aos debates para emitir a decisão, devendo fazer parte da mesa julgadora, pelo menos, um membro da região onde o fato ocorreu. Todas as decisões passarão por um reexame com todos os membros do

Comitê e, caso de aprovação, será publicada no *site* oficial. Não sendo deferido o pedido de liberação do conteúdo pela maioria, o Comitê deverá convocar um novo painel para reanalisar o caso. (Oversight Board)

Pode-se observar que o modelo escolhido para o fluxo de julgamento do Comitê de Supervisão é muito semelhante aos julgamentos de Tribunais e Cortes Superiores. Tendo uma limitação de tema de relevância significativa, como também, a discussão em plenário, em caso de reanálise para pedidos indeferidos.

Ainda, assim, como os ordenamentos jurídicos em geral, a decisão tomada pelo Comitê deve ser fundamentada, obedecendo alguns requisitos essenciais, tais quais: (1) um breve relato do caso; (2) a autoridade e escopo, sendo descritas as bases legais do estatuto do comitê para a competência do julgamento; (3) padrões relevantes, no qual serão elencados os dispositivos os quais o comitê baseará suas decisões através das políticas de conteúdo do *Facebook*, dos valores do *Facebook* e dos padrões de direitos humanos descritos nas leis internacionais que basearam o julgamento; (4) a declaração do criador do conteúdo, momento em que é dado o direito ao contraditório ao usuário da postagem removida das plataformas; (5) a explicação da plataforma sobre o porquê decidiu pela remoção do conteúdo da rede; (6) o envio de terceiros, meio pelo qual serão analisados os comentários públicos relacionados no caso em análise³⁹; (7) a análise do Comitê de Supervisão em conformidade com as políticas de conteúdo, os valores do *Facebook* e as responsabilidades sobre direitos humanos do *Facebook*⁴⁰. Por fim, o Comitê emitirá sua decisão e publicará, também, uma declaração de orientação para políticas futuras às plataformas. (Oversight Board)

Após 3 anos de funcionamento, em junho de 2023 o *Oversight Board* publicou seu primeiro relatório sobre o trabalho do Comitê, no qual discorre sobre os primeiros pareceres e o impacto que suas recomendações estão sendo aplicadas e recebidas pelas plataformas do *Meta Inc.* e por seus usuários. (Oversight Board, 2023)

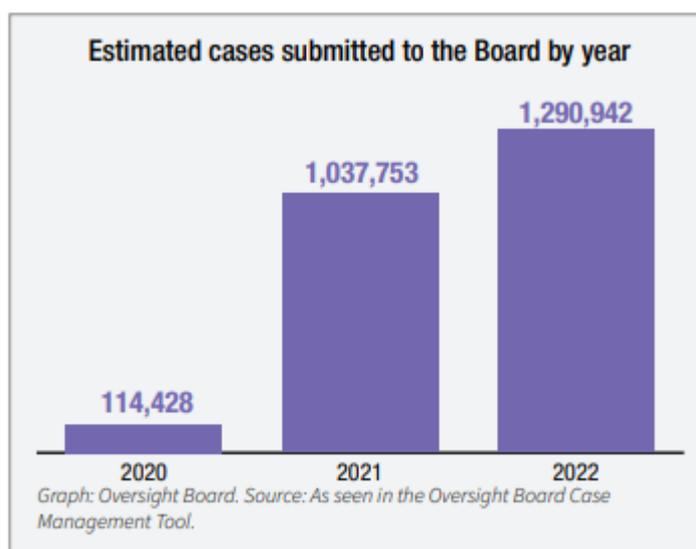
³⁹ Quando o Comitê escolhe os casos para julgamento, ele publica em sua página – aba “Notícias” – e abre espaço de 10 dias para que especialistas e grupos/organizações interessados possam contribuir, através de suas perspectivas diversificadas, para a análise e julgamento de cada caso. São fornecidos alguns questionamentos para ajudar a guiar os comentários públicos. (Oversight Board)

⁴⁰ A parte relacionada à responsabilidade sobre os direitos humanos pelo *Facebook* contemplará critérios semelhantes àqueles dispostos no documento da Assembleia Geral da ONU, de 2019, abordado no capítulo anterior, acerca da legalidade (clareza e acessibilidade das regras), o objetivo legítimo, bem como a proporcionalidade e necessidade (analisado o contexto, o status do discursista, intenção, conteúdo e formato, extensão e alcance, iminência do dano). (Oversight Board)

Segundo o relatório em 2022 o Comitê recebeu cerca de 1,3 milhões de casos de usuários para análise, um quarto a mais que em 2021. Ao todo foram 50 casos selecionados e 32 decisões originárias revertidas da plataforma *Facebook*. Além disso, desde janeiro de 2021 até o início de abril de 2023, o *Oversight Board* havia realizado cerca de 191 recomendações ao *Meta Inc.*, tendo dois terços dessas recomendações sido implementadas pela plataforma, seja de forma parcial ou total. (Oversight Board, 2023)

Conforme pode ser observado na figura abaixo, houve um aumento significativo a submissão dos casos ao Comitê desde o seu funcionamento:

Figura 1 – Estimativa de submissão de casos ao Comitê por ano



Fonte: *Oversight Board* (2023)

As principais recomendações do Comitê ao *Meta* foram em relação: a informação aos usuários do motivo pelo qual o conteúdo estava sendo removido, dessa forma as plataformas *Facebook* e *Instagram* passaram a publicar o motivo da remoção da postagem, qual a política da comunidade foi violada e, ainda, se a remoção ocorreu por análise humana ou automatizada; transparência e consistência nas regras; a proteção da livre manifestação dos profissionais de jornalismo, manifestantes e ativistas. (Oversight Board, 2023)

Nota-se que desde o funcionamento do Comitê a submissão de casos ao seu crivo está crescendo cada vez mais. O que, em parte, demonstra que seu desempenho está se mostrando confiável por meio dos usuários das plataformas, às quais o *Board* possui competência para julgamento de casos. Da mesma forma me

que o *Meta Inc.*, aponta estar em harmonia com a autoridade de seu “órgão julgador”, buscando se adaptar e implementar as recomendações realizadas por meio de suas decisões, o que demonstra que o *Oversight Board* está cumprindo o papel para o qual foi designado, de forma independente e transparente.

Assim, no *Annual Report* de 2022, o Comitê conclui sobre a grande responsabilidade que possui perante o auxílio na moderação de conteúdo, almejando a expansão de suas atribuições dentro do ciberespaço:

At its best, social media can be an unparalleled catalyst for global connection and conversation. To reap these benefits, while containing its harms, is a daunting task. As a pioneering entity, the Oversight Board continues to apply lessons learned to improve itself. We believe that social media companies make content moderation decisions in a fairer, more principled way if they base those decisions on international human rights standards. We stand ready to share what we have learned so far with other companies and organizations that share our goals of increasing transparency and improving how people are treated online. (Oversight Board, 2023, p. 64)⁴¹

Entendendo-se, assim, o funcionamento e a atribuição que possui o *Oversight Board*, a presente pesquisa iniciará análise da suspensão de contas de Donald Trump, após a derrota nas eleições presidenciais em 2020, iniciando com a contextualização do ocorrido no fatídico 6 de janeiro e, posteriormente, o julgamento que manteve a suspensão da conta do ex-presidente.

4.2 Relato sobre os atos que antecederam a suspensão das contas de Donald Trump

Para compreender melhor o desfecho que levou à suspensão das contas do candidato à Presidência dos Estados Unidos, Donald Trump, é necessário trazer alguns pontos estruturais que culminaram na situação que será analisada no último tópico.

É necessário contextualizar, portanto, que polêmicas envolvendo o ex-

⁴¹ Na melhor das hipóteses, as redes sociais podem ser um catalisador incomparável para conexões e conversas globais. Colher estes benefícios, ao mesmo tempo que contém os seus danos, é uma tarefa difícil. Como entidade pioneira, o Conselho de Supervisão continua a aplicar as lições aprendidas para melhorar. Acreditamos que as empresas de redes sociais tomam decisões de moderação de conteúdo de uma forma mais justa e baseada em princípios se basearem essas decisões em padrões internacionais de direitos humanos. Estamos prontos para partilhar o que aprendemos até agora com outras empresas e organizações que partilham os nossos objetivos de aumentar a transparência e melhorar a forma como as pessoas são tratadas online. (Conselho de Supervisão, 2023, p. 64, tradução livre)

presidente dos Estados Unidos nas redes sociais já vêm ocorrendo desde a corrida presidencial de 2016, no qual culminou no escândalo *Cambridge Analytica x Facebook*.

A primeira campanha eleitoral de Trump à presidência se iniciou com um movimento de baixo custo do candidato Republicano nas eleições presidenciais de 2016 nos Estados Unidos, o que não se sabia é que por trás havia um grandioso contrato com a empresa de consultoria *Cambridge Analytica*, e que acabou se aproveitando das ferramentas contidas na rede social do *Facebook* para distribuir e disseminar falsas notícias e discursos de ódios sobre aqueles grupos que seus eleitores entendiam como “inimigos” da nação⁴². (Empoli, 2019)

A referida empresa de consultoria britânica, a qual foi contratada pela campanha de Trump, se utilizou de um aplicativo de enquete *thisisyourdigitallife* e colheu, além dos dados de quem realizava o *quiz*, os dados de seus amigos na rede social do *Facebook*. De acordo com Empoli (2019), cerca de 270 mil pessoas fizeram o teste, tendo sido coletados os dados de, pelo menos, 50 milhões de usuários, os quais serviram para traçar um perfil da comunidade, criar grupos de acordo com suas características e preferências, para assim direcionar o conteúdo específico para cada usuário, como exibição de anúncios com fins políticos.

É evidente que a utilização das mídias sociais e da análise de seus dados para direcionamento de conteúdo político e ideológico, muitos contendo discurso de ódio e *fake news*, não é e nem foi uma invenção de Trump, mas não se pode negar que a sua estratégia de campanha deu certo, tanto é que foi eleito para a presidência dos Estados Unidos em 2016.

E aqui vale ressaltar e relembrar o papel amplificador que as mídias sociais possibilitaram para a disseminação desse tipo de conteúdo, pois a quantidade de informações que recebemos a todo instante é massificada que chega ao ponto em que os usuários não passam a não mais distinguir o que é verdade ou mentira. Assim, não importando o assunto ou tópico discutido, as dúvidas sempre serão levantadas e

⁴² A ideia de trazer a narrativa de “inimigos da nação” é um mecanismo bastante utilizado no discurso de ódio, sendo facilmente identificável em diversos períodos históricos mundiais. Pereira (2017, p. 168) compreende que a criação desse discurso é intencional, como forma de manter grupos ou elites dominantes no comando da estrutura de poder, nutrindo a ideia de que existe um inimigo a ser combatido. Dessa forma, ao eleger quem fará o papel de inimigo, passa-se a rotulá-lo e recriminá-lo por meios de opressão, para que a comunidade passe a compreendê-los como uma ameaça ao seu livre arbítrio e a sua própria segurança. (Schäfer; Leivas; Santos, 2015, p. 147; Sémelin, 2009, p. 25-41)

isso viabiliza a disseminação de narrativas manipuladoras com muito mais facilidade. (O dilema..., 2020)

Essa movimentação de manipulação, muito em alta no ciberespaço, foi identificada e descrita como a Era da Pós-verdade⁴³, na qual o mundo modificou de forma significativa o embate entre os campos políticos e intelectuais, uma vez que sua inflexibilidade e descredibilização das “instituições democráticas estão sendo abaladas em suas bases por uma onda de populismo ameaçador”. A partir daí, instaura-se uma dinâmica do desmoronamento “do valor da verdade”, em que a verdade dos fatos não é mais uma moeda de troca dentro do meio político. (D’Ancona, 2018, p. 19-20)

Segundo D’Ancona (2018, p. 34), na Era da Pós-verdade, o cidadão começa a recusar a verdade dos fatos, deixando-se levar por suas emoções, convicções e ideologias na formação de opinião. Não importa o quanto os fatos estejam à frente de determinado grupo social ou ideológico, ele só irá conseguir acreditar naquilo irá impactar seu íntimo.

Empoli (2019), trouxe em sua obra o impacto que a utilização de conteúdos de *fake news* e discurso de ódio tem em relação à crescente disseminação de compartilhamento nas redes sociais. Segundo estudo do MIT (Massachusetts Institute of Technology) as notícias falsas têm 70% mais probabilidade de serem retuitadas no *Twitter* do que a verdade e alcançam suas primeiras 1.500 pessoas seis vezes mais rápido, esse efeito possui uma maior repercussão quando se trata de notícias políticas do que com outras categorias. Durante a pesquisa, foi observado que enquanto os *bots* espalham informações verdadeiras e falsas nas mesmas taxas, os usuários se sentiam mais atraídos por informações novas e incomuns, como as *fakes news* costumam ser.

Esses fenômenos que foram aperfeiçoados na Era Digital são bastante utilizados no campo político e Donald Trump coleciona conteúdos de desinformação postados por ele próprio ou por sua equipe em suas redes sociais, dentre os quais pode-se de destacar as falsas alegações de que Obama não é americano, que o México enviava estupradores e criminosos para os Estados Unidos, que os

⁴³ A palavra pós-verdade foi eleita como a palavra do ano, em 2016, pelo Dicionário Oxford sendo classificada como algo “relativo a ou que denota circunstâncias nas quais fatos objetivos são menos influenciadores na formação da opinião pública do que apelos à emoção ou à crença pessoal”. Entretanto, o que discerne a mentira comum na área política e a mentira na Era da Pós-verdade é a resposta da sociedade perante ao argumento trazido. (Mcintyre, 2018, p. 1; D’Ancona, 2018, p. 34)

Democratas criaram o Estado Islâmico (tendo Obama e Hillary Clinton como cofundadores), que crianças eram quase que imunes à Covid-19, dentre outros. Em uma reportagem publicada pela CNBC, foi constatado que das 100 postagens mais populares de Trump no *Twitter*, 36 continham falsidades relacionadas à eleição de 2020. (Modelli, 2020; Rattner, 2021)

Como pode se notar, Trump passou a campanha de sua reeleição utilizando-se de fake news para desestabilizar não somente seus adversários, mas criou um clima de instabilidade e desconfiança política de seus eleitores para com o processo eleitoral daquele ano. E isso ficou cada vez mais evidente após o início da contagem de votos após o encerramento das votações no país, no dia 03 de novembro de 2020.

Como o início da apuração dos votos e com a constatação que o resultado seria acirrado entre os candidatos do partido Democrata e do partido Republicado, no dia 05 de novembro, Donald Trump se declarou o vencedor das eleições e iniciou novos ataques ao sistema eleitoral norte-americano, o qual resultou popular *tweet* “STOP THE COUNT”:

Figura 2 – *Tweet* de Trump “*Stop the count*”

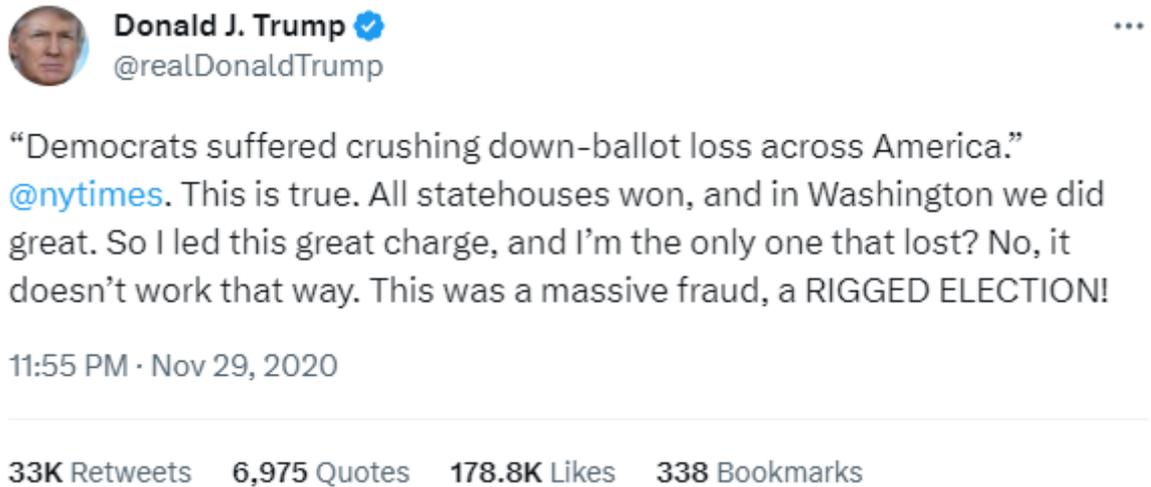


Fonte: *Twitter* (2020).

Segundo Trump, até aquele momento ele havia vencido as eleições de 2020 e se tornado o 46º presidente dos Estados Unidos. Todos os votos que estavam sendo contabilizados a partir de então era mero produto de fraude dos Democratas, com intuito de derrubar o candidato republicado da presidência. Assim, prosseguiu-se

com uma série de ataques e acusações de fraude eleitoral ao longo das semanas até o dia 06 de janeiro de 2021:

Figura 3 – *Tweet* de Trump afirmando fraude nas eleições



Fonte: *Twitter* (2020).

Figura 4 – *Tweet* de Trump afirmando que havia vencido as eleições

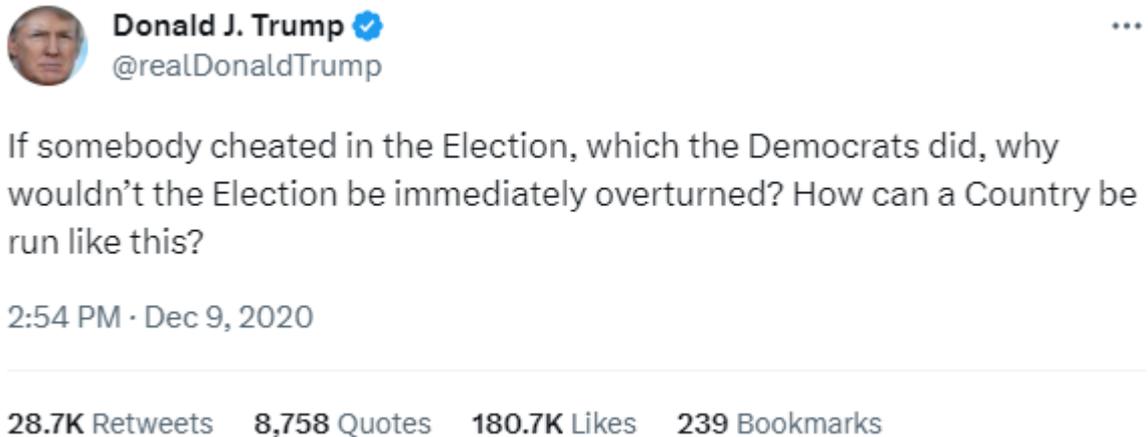


Fonte: *Twitter* (2020).

Com as postagens, não somente no *Twitter*, mas também nas redes sociais do *Facebook* e *Instagram*, Trump alimentava seus eleitores com desinformações sobre a contagem de votos, diversas foram as postagens escritas, vídeos e a

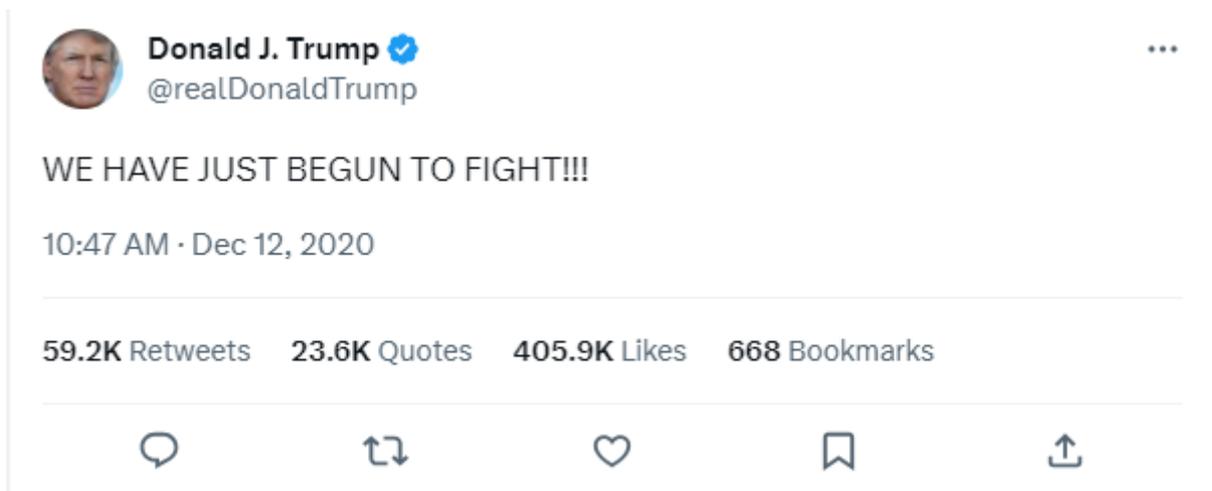
enganosa comprovaç o de que a eleiç o havia sido roubada deles. Sim, para Trump e seus eleitores, a democracia dos Estados Unidos estava se corrompendo:

Figura 5 – *Tweet* de Trump afirmando que houve trapaça nas eleiç es



Fonte: *Twitter* (2020).

Figura 6 – *Tweet* de Trump informando que a “batalha” estava comeando



Fonte: *Twitter* (2020).

Figura 7 – *Tweet* de Trump afirmando corrupção nas eleições



Fonte: *Twitter* (2020)

Ao longo que a posse de John Biden ia se aproximando, o candidato derrotado, Donald Trump, chamou seus simpatizantes e eleitores para protestar na capital estadunidense contra o resultado das eleições e marcou na data em que ocorreria a certificação da eleição pelo Congresso:

Figura 8 - *Tweet* de Trump chamando seus eleitores para a capital no dia 06/01



Fonte: *Twitter* (2020).

Figura 9 – Tweet de Trump convocando seus eleitores para o protesto em 06/01



Fonte: Twitter (2020).

No dia 06 de janeiro de 2021, durante o comício promovido por Trump, insuflou seus seguidores a partirem em direção ao Capitólio para terem suas “vozes” ouvidas pelo Congresso, visto que os congressistas estavam reunidos na casa legislativa. Trump, então se dirigiu aos eleitores dizendo: *"I know that everyone here will soon be marching over to the Capitol building to peacefully and patriotically make your voices heard"*⁴⁴. (Capitol riots..., 2021)

Há quem defenda que Trump jamais quis que a caminhada ao Capitólio resultasse no ataque violento do dia 6 de janeiro, mas há quem afirme que os discursos inflamados do ex-presidente contribuíram diretamente para que seus eleitores tomassem as vias de fato naquele dia. De certa forma, querendo ou não este desfecho, as palavras de Trump ecoaram e instigaram seus seguidores de quem precisavam lutar com suas próprias mãos para defender a democracia estadunidense, visto que estava sendo corrompida com a eleição mais fraudulenta da história do país.

Assim, durante o processo oficial de confirmação da vitória de Biden, o Congresso foi invadido por grupos de eleitores de Donald Trump, muitos dos quais se encontravam armados. Os congressistas e funcionários foram retirados às pressas, enquanto os invasores depredavam janelas, portas e móveis, quebravam vidraças em busca de uma vingança em nome da “democracia”. Com pouco efetivo policial trabalhando naquela tarde no Capitólio, o caos se instaura.

⁴⁴ “Sei que todos aqui irão em breve marchar até ao edifício do Capitólio para fazer ouvir as suas vozes de forma pacífica e patriótica”. (Capitol riots..., 2021)

Horas após a invasão, Donald Trump publica em seu perfil oficial no *Instagram* e no *Facebook* um vídeo com uma mensagem aos invasores, a qual está transcrita abaixo:

*I know your pain, I know you're hurt, We had an election that was stolen from us. It was a landslide election and everyone knows it, especially the other side. But you have to go home now. We have to have peace. We have to have law and order. We have to respect our great people in law and order. We don't want any hurt. It's a very tough period of time. There's never been a time like this, where such a thing happened, where they could take it away from all of us. From me, from you, from our country. This was a fraudulent election, but we can't play into the hands of these people. We have to have peace. So go home. We love you, you're very special. We've seen what happens, you see the way others are treated that are so bad and so evil. I know you how feel. But go home, and go home in peace.*⁴⁵ (Capitol riots..., 2021)

Como pode ser observado, na mensagem o ex-presidente se compadecia com a dor dos invasores, buscando uma justificativa para o injustificável, volta a afirmar que a eleição foi fraudada e que esses “fraudadores” haviam tirado tudo deles. Ainda, para finalizar, ele enaltece seus seguidores, falando o quanto os ama e que sabe como eles estão se sentindo naquele momento. Além disso, Trump dá a entender que a repressão em cima dos invasores foi arbitrária e desproporcional.

O *Facebook* removeu o vídeo, mas mesmo após a remoção, o ex-presidente voltou a postar conteúdo, agora na forma escrita, enaltecendo aqueles que invadiram o Capitólio como grandes patriotas. A plataforma, novamente, removeu o conteúdo e bloqueou as contas do *Facebook* e *Instagram* de Trump, impedindo-o de publicar por 24h.

Naquele dia cinco pessoas morreram em decorrência da invasão, dentre civis e policiais, cento e quarenta policiais ficaram feridos. Outros quatro agentes de segurança cometeram suicídio nos meses seguintes. (Tortella, 2022)

Os danos materiais e psicológicos de quem viveu aquele momento são imensos, além dos danos causados à democracia estadunidense e ao processo

⁴⁵ Eu conheço a sua dor, eu sei que você está ferido. Tivemos uma eleição que nos foi roubada. Foi uma eleição esmagadora e todos sabem disso, especialmente o outro lado. Mas você tem que ir para casa agora. Temos que ter paz. Temos que ter lei e ordem. Temos que respeitar nosso grande povo na lei e na ordem. Não queremos nenhuma dor. É um período muito difícil. Nunca houve um momento como este, onde tal coisa aconteceu, onde eles pudessem tirar isso de todos nós. De mim, de você, do nosso país. Esta foi uma eleição fraudulenta, mas não podemos fazer o jogo destas pessoas. Temos que ter paz. Então vá para casa. Nós te amamos, você é muito especial. Vimos o que acontece, você vê a maneira como os outros são tratados, que são tão maus e tão perversos. Eu sei como você se sente. Mas vá para casa e volte para casa em paz. (Capitol riots..., 2021)

eleitoral. A suspensão das contas de Donald Trump foi uma reação rápida da plataforma para amenizar os danos ocorridos e conter danos futuros. Após a suspensão de 24 horas, a rede social estendeu por mais algumas semanas a medida e após encaminhou o caso para o Comitê de Supervisão analisar se sua conduta estava correta.

Assim, passa-se ao último tópico da pesquisa, no qual irá analisar o julgamento da suspensão de contas do ex-presidente dos Estados Unidos, Donald Trump, através das premissas e conceitos estudados sobre liberdade de expressão, seus limites e o funcionamento do *Oversight Board*.

4.3 A análise do julgamento do caso 2021-001-FB-FBR

Inicialmente, conforme foi abordado no subcapítulo anterior, após o ex-presidente Donald Trump publicar em suas contas no *Facebook* e *Instagram* vídeo alegando que as eleições foram fraudadas, que entendia o sofrimento e a mágoa daqueles que estavam revoltados com essa situação, mas que eles deveriam retornar para suas casas e ficarem seguros, pois ele os amava muito e reforçando que a reação contra os invasores do Capitólio foi desproporcional para com eles, as contas de Trump foram suspensas pelo período de 24h das plataformas do *Meta Inc*.

No dia seguinte, a plataforma fez uma análise adicional das publicações de Trump, contanto com informações de outras redes sociais e da gravidade do atentado do dia anterior e estendeu o bloqueio por tempo indefinido e, pelo menos, por 2 semanas até a transferência de poder na presidência.

Após a posse de John Biden, o *Facebook* anunciou que havia encaminhado o caso para o Comitê, questionando aos membros se havia tomado a decisão correta no momento do bloqueio e solicitou recomendações sobre a suspensão para líderes políticos.

O caso foi remetido ao Comitê que passou à análise do caso, por meio de todo o processo já elucidado anteriormente.

Para os julgadores do *Oversight Board*, as publicações do dia 06/01/2021 violaram gravemente os Padrões da Comunidade do *Facebook* e as Diretrizes do *Instagram* que proíbem exaltar ou apoiar pessoas envolvidas em violência. E que ao manter a narrativa, sem fundamento algum, de que houve fraude eleitoral e as chamadas de apoio de seus eleitores, o ex-presidente fomentou um risco grave de

possíveis atos de violência e suas palavras de apoio aos envolvidos legitimou a ação violenta dessas pessoas. Além de toda a influência que Trump tinha e o grande alcance de suas publicações. (Oversight Board, 2021)

Ademais, o comitê entendeu ter sido justificada a suspensão das contas de Trump de 6 a 7/01/2021, mas que não foi apropriada a imposição, pela plataforma, de uma suspensão indefinida, sem critérios em relação a quando e se a conta será restaurada, pois isso violaria as regras e penalidades da própria rede, que impõe a suspensão de contas com prazo determinado ou a desativação permanente. (Oversight Board, 2021)

Na decisão, há a advertência à plataforma de que quem deve criar penalidades aplicáveis e adequadas é ela mesma, sendo atribuição do comitê apenas a garantia de que as decisões e penalidades aplicadas pela rede condizem com suas regras/políticas internas e se estas estão de acordo com as prerrogativas dos direitos humanos. E o *Facebook* ao aplicar sanções ao usuário de forma vaga e fora dos padrões da comunidade e requerer que o Comitê resolva, é se esquivar de sua própria responsabilidade, aplicando e justificando uma penalidade já definida em seus padrões. (Oversight Board, 2021)

Em relação ao pedido do *Facebook* de solicitar recomendações sobre a suspensão de líderes políticos, o Comitê entendeu que não seria benéfico fazer uma distinção inflexível de líderes políticos e outros usuários influentes, visto que os últimos também podem contribuir para causas com risco de graves danos. Seria necessário, portanto, uma análise no caso concreto de probabilidade e iminência de dano, ou seja, quando um usuário influente faz uma publicação que representa uma grande probabilidade de danos iminentes, a plataforma deverá agir imediatamente e aplicar suas regras, inclusive quando se tratar de conteúdo de valor jornalístico⁴⁶, não devendo estas serem priorizadas quando for necessária uma medida urgente para impedir um dano significativo. (Oversight Board, 2021)

Na decisão foi sugerido ao *Facebook* que quando precisar suspender uma conta de um usuário influente, deverá explicar publicamente as regras no momento da imposição da sanção, devendo estas [regras] garantirem que o tempo de suspensão possa reduzir o risco de dano significativo, cabendo a plataforma

⁴⁶ Segundo o Comitê, o conteúdo de valor jornalístico diz respeito à permissão da plataforma em manter o conteúdo no ar (mesmo tendo violado as regras/políticas da comunidade) caso seja considerado como importante e de interesse público. (Oversight Board, 2021)

reanalisar os riscos antes do término da suspensão. Caso seja identificada, ainda, a probabilidade e iminência de dano, outra suspensão poderá ser imposta até a quantidade de medidas necessárias para proteger a segurança pública, desde que seja proporcional ao risco. (Oversight Board, 2021)

Por fim, o Comitê advertiu que chefes de estado e funcionários de alto escalão dos governos podem ter uma capacidade maior de causar dano do que outras pessoas. Caso essas pessoas publiquem repetidamente mensagens que representem um risco iminente de dano, o *Facebook* deverá suspender a conta por período suficiente para a proteção de eventuais prejuízos, aplicando um período de suspensão suficientemente longo para deter a conduta indevida, podendo acarretar na exclusão da conta ou da página. (Oversight Board, 2021)

Diante do que foi estudado até o momento, pode-se constatar que o Comitê de Supervisão observou se houve a devida aplicação dos Padrões da Comunidade do *Meta*, quando analisou as implicações e regramentos da suspensão de contas da plataforma, demonstrando que ao ter suspenso por tempo indeterminado as contas de Donald Trump das redes sociais *Facebook* e *Instagram*, o *Meta* estaria violando sua própria regulamentação interna. Devendo assim, ao realizar a suspensão da conta de Trump, determinar o tempo pelo qual perdurará a sanção.

No que tange a aplicação dos limites à liberdade de expressão, pode-se observar que o Comitê se utilizou de parâmetros internacionais para a compreensão e julgamento do caso, empregando os mesmos critérios estabelecidos tanto pela ONU quanto pela OEA para a imposição de limites à liberdade de expressão, os quais são: legalidade, legitimidade, necessidade e proporcionalidade. Estes critérios buscam auxiliar em uma melhor análise à medidas extremas e complexas como a limitação da livre manifestação.

Através desses critérios, foi possível verificar se havia regramentos nos termos de uso da plataforma para a imposição de moderação de determinados conteúdos ou de sanções de suspensão ou exclusão de contas em razão de postagem violadora de critérios determinados previamente pela plataforma, por meio do princípio da legalidade, o que de fato ocorreu, pois segundo observado pelos julgadores, a plataforma aplicou bem a adequação na categoria “Indivíduos e Organizações Perigosas”, tendo em vista que Trump promoveu elogios e apoiou os invasores do Capitólio. Identificado, portanto, a violação dos padrões da comunidade.

Cumprido ressaltar que a suspensão por tempo indeterminado não consta

nos critérios de suspensão de usuário e, por isso, o Comitê pontuou que a plataforma deveria ter estabelecido um prazo determinado para a duração da sanção.

Ademais, percebe-se que o Comitê de Supervisão utiliza-se, também, a legislação e os entendimentos acerca da liberdade de expressão abrangidos pelos padrões internacionais de direitos humanos, tais como a declaração universal de direitos humanos, pacto internacional de direitos civis e políticos, convenção internacional sobre eliminação de todas as formas de discriminação racial, plano de ação de Rabat. Padrões estes que também estão incorporados pelo *Meta Inc.* e adotados na elaboração dos padrões da comunidade.

Vale o destaque aqui para a defesa de Donald Trump que solicitou a aplicação da legislação norte-americana para a interpretação dos critérios de legalidade, legitimidade, necessidade e proporcionalidade, uma vez que se estava realizando um julgamento de um líder de Estado e, portanto, os princípios jurídicos deveriam se adequar conforme a localidade em que o líder está ou estava governando. (Oversight Board, 2021)

O pedido, claro, foi negado pelo *Board*, esclarecendo na decisão que a *First Amendment* não é a legislação aplicável para controle de empresas privadas e que o princípio fundamental à liberdade de expressão também está previsto no texto da PIDCP, mas que não se pode olvidar quando um agente do Estado viola diretrizes internacionais de direitos humanos, as quais deveria salvaguardar. (Oversight Board, 2021)

Ainda, os julgadores sustentaram que mesmo que a legislação internacional autorize a limitação à liberdade de expressão, esta restrição deve cumprir alguns requisitos:

Qualquer restrição deve cumprir três requisitos: as regras devem ser claras e acessíveis, deve ter um objetivo legítimo e deve ser necessária e proporcional ao risco de dano. O Comitê usa esse teste de três partes para analisar as ações do Facebook quando a empresa restringe conteúdo ou contas. Os princípios da Primeira Emenda nos termos da lei dos EUA também insistem que as restrições sobre a liberdade de expressão impostas por ação do Estado não podem ser vagas, mas devem ter motivos governamentais importantes e devem estar estreitamente adaptadas ao risco de dano. (Oversight Board, 2021)

Assim, pode ser observado que o Comitê também analisa critérios do risco de dano causado pelo discurso, muito semelhante à aplicação da jurisprudência norte-americana do teste de *Brandenburg*, no qual busca observar se conteúdo postado

teria o condão de causar risco iminente de um ato ilícito, mas talvez buscando trazer parâmetros tão rígidos da neutralidade do discurso quanto ao sistema estadunidense.

Quanto ao critério da legitimidade, foi analisada se a rede social justificou de forma clara as razões pelas quais está aplicando a sanção ao usuário e se essa legitimação está em conformidade com os *standarts* de direitos humanos, não para censurar o discurso por simples antipatia a este, mas sim como forma de evitar ou reduzir danos causados por ele, conforme consta no texto transcrito do julgado:

O requisito de objetivo legítimo significa que qualquer restrição à liberdade de expressão deve ser por uma finalidade indicada no Artigo 19, parágrafo 3 do PIDCP e esta lista de objetivos é exaustiva. Os objetivos legítimos incluem a proteção da ordem pública, bem como o respeito pelos direitos de terceiros, incluindo os direitos à vida e à segurança e de participar de eleições e ter o resultado respeitado e implementado. Um objetivo não seria legítimo se usado como pretexto para tolher a liberdade de expressão; por exemplo, mencionar os objetivos de proteção da segurança ou dos direitos de terceiros para censurar o discurso simplesmente por ele ser desagradável ou ofensivo (Comentário Geral nº 34, parágrafos 11, 30, 46 e 48). A política do Facebook sobre exaltação e apoio de indivíduos envolvidos em “eventos violadores”, atos de violência ou atividade criminal estava de acordo com os objetivos já citados. (Oversight Board, 2021)

E em relação à necessidade e proporcionalidade, o *Oversight Board* entendeu que a medida foi imprescindível para a contenção dos danos presentes e, também, em parte proporcional em razão da manutenção da suspensão da conta, como forma de evitar novos danos que poderiam vir a acontecer, a partir de novas manifestações do ex-presidente em seus perfis nas redes sociais. Entretanto, deixou de observar a proporcionalidade ao suspender o perfil por tempo indeterminado:

Na análise desses fatores, o Comitê conclui que a violação neste caso foi grave em termos de danos aos direitos humanos. A imposição do Facebook de restrição da conta, em 6 de janeiro, e extensão dessa restrição, em 7 de janeiro, foi necessária e adequada. [...] Para a minoria, essa análise mais ampla seria crucial para fundamentar a avaliação do Facebook de uma penalidade adequada ao ocorrido em 7 de janeiro, o que serviria como impedimento para outros líderes políticos e, onde apropriado, como uma oportunidade de reabilitação. Além disso, se o Facebook optasse por impor uma suspensão com prazo determinado, a análise de risco necessária para reintegração também levaria em conta esses fatores. (Oversight Board, 2021)

Encerra-se, portanto, a análise do julgamento da suspensão das contas do ex-presidente Donald Trump, tendo sido analisada em consonância ao que foi estudado ao longo desta pesquisa. Assim, passa-se a conclusão a presente dissertação.

CONCLUSÃO

Conforme pontuado na introdução dessa pesquisa, o objetivo principal é compreender de que maneira o *Oversight Board* pode garantir a proteção do direito humano à liberdade de expressão, a partir de um estudo dos Padrões da Comunidade do *Facebook*, os *standarts* Internacionais de direitos humanos sobre liberdade de expressão e a legislação norte-americana, tendo em conta a análise do caso referente à suspensão das contas do *Facebook* e *Instagram* do ex-presidente dos Estados Unidos, Donald Trump. Para isso, o estudo perpassou por diversas áreas para buscar encontrar uma resposta à problemática trazida.

Sendo assim, o primeiro capítulo tratou de descrever as revoluções da humanidade desde quando o ser humano passou a se agrupar e formar suas comunidades e cidades, a partir disso a organização da sociedade foi se modificando ao longo do tempo até que chegaram às revoluções industriais. Essas revoluções, sem dúvidas, trouxeram uma nova dinâmica social e foi um movimento incipiente para que novas tecnologias pudessem surgir.

Neste capítulo inicial, também, tratou-se de relatar toda a transformação que as revoluções do conhecimento e tecnológicas trouxeram para a sociedade, as quais modificaram a forma de interpretar e tratar a comunicação, passando a ter um valor mercadológico com o surgimento da sociedade da informação e comunicação e da sociedade em rede.

Inegável dizer que o ciberespaço e a ascensão da comunicação por meio das mídias/redes sociais levaram o ser humano para uma nova experiência: a vida *onlife*. Isso trouxe um grande impacto para vida ao redor do mundo, modificando o mercado, a economia, a comunicação e até a forma de se relacionar. O que culminou para a discussão levantada no último tópico do primeiro capítulo, que buscou trazer ao debate a necessidade do reconhecimento dos novos direitos que a Era Digital impôs ao mundo, de como podemos proteger a sociedade da tecnologia, que está sempre suscetível a manipulações e carentes de uma regulamentação para a proteção daqueles que as utilizam.

O segundo capítulo buscou compreender a aplicação do direito à liberdade de expressão e identificar seus limites legais por meio dos padrões internacionais de direitos humanos, padrões da comunidade do *Facebook* e a legislação norte-americana.

Neste capítulo foi possível observar a universalidade do que se entende por liberdade de expressão, mas, também, que a imposição de critérios limitadores à livre manifestação pode ser distinta, dependendo do ordenamento jurídico em que é aplicado. Enquanto o *free speech* possui uma maior proteção na legislação e jurisprudência norte-americana, o direito internacional abarca diversos critérios possíveis para a moderação do discurso, buscando requisitos, a partir da interpretação dos direitos humanos, para contrapesar o conflito de direitos fundamentais, os quais devem ser resguardados, mas não podem ser sobrepujados para violar o direito de outrem.

Foi abordado, também, a questão da moderação de conteúdo nas redes sociais aplicada ao direito internacional, a qual, por muito tempo, careceu de uma regulamentação específica, tendo em vista que quando se passou a visualizar os impactos que as violações de direitos humanos estavam causando para além do ciberespaço, o problema passou a ser de interesse público. Voltando-se, assim, a atenção do poder público e da sociedade civil para a observância e complementação dessa lacuna jurídica.

Ainda, fez-se necessário apontar como funciona a autorregulação dos provedores intermediários, para poder melhor entender como surgiram e como se aperfeiçoaram os critérios estabelecidos pelo *Meta Inc.* para a definição dos termos de uso pelos usuários das plataformas da empresa.

O terceiro e último capítulo, teve como objetivo específico analisar o julgamento da suspensão de contas do ex-presidente dos Estados Unidos, Donald Trump, através das premissas e conceitos estudados sobre liberdade de expressão, seus limites e o funcionamento do *Oversight Board*.

Para isso, explicou-se a finalidade do Comitê de Supervisão, a forma e os critérios que são adotados para a escolha e julgamento dos casos e de que forma é fundamentada suas decisões, as quais se valem dos *standarts* internacionais de direitos humanos para solucionar os casos que devem ser deliberados.

Contextualizar-se, também, como foi a crescente influência de Donald Trump e sua ascensão à presidência dos Estados Unidos por meio da utilização de redes sociais para alcançar um público bastante devoto e extremista, sendo bastante utilizadas técnicas de disseminação de desinformação, discurso de ódio, dentre outros. Toda essa movimentação realizada durante anos, colaborou para os atos que aconteceram durante a campanha de Trump à reeleição e aos atos de extrema

violência ocorridos no dia 06/01/2021, que culminou com a invasão ao Capitólio e a morte de cinco pessoas durante o ocorrido.

O *Oversight Board* foi uma iniciativa inovadora de Mark Zuckerberg. Que, de certa forma, buscou mitigar os danos que a manipulação algorítmica e a disseminação de conteúdos violadores vêm ocorrendo e chamando a atenção do interesse público e da sociedade civil desde a divulgação do escândalo da *Cambridge Analytica*.

O também chamado “Tribunal do *Facebook*” vem atuando e julgando casos de grande repercussão nas redes sociais do *Facebook* e *Instagram*, e até o momento se vislumbrou a independência, seriedade e transparência do Comitê perante a condução da análise e decisões acerca dos casos que lhes são apresentados.

A escolha para a aplicação dos padrões internacionais de direitos humanos aparenta ser uma decisão acertada para buscar contrapesar a escolha da salvaguarda do direito fundamental à liberdade de expressão e a possibilidade de limitação do *free speech* nas redes sociais.

O caso de Trump demonstrou que o escopo do Comitê está bastante estruturado e que não permite a implementação de normas que não transigem com os padrões internacionais de direitos humanos, conforme se pode observar no caso de Trump, em que foi solicitada a aplicação da Primeira Emenda para julgamento do caso, mas não acolhido o argumento em razão da limitação do regramento norte-americano.

Ademais, nota-se que mesmo havendo uma gradativa e significativa mudança nos regulamentos internos do *Meta Inc.* ao longo dos anos, os quais foram cada vez mais adaptados para incorporar as regulamentações internacionais de direitos humanos, as plataformas ainda possuem a prerrogativa de autorregulação. Ou seja, ainda depende de qualquer plataforma a intenção de acatar ou não essas diretrizes.

Justo por isso que há um interesse público na regulamentação internacional para moderação de conteúdo e se procura, já há algum tempo, meios de concretizar essa ideia. Levanta-se a solução por meio do chamado “Constitucionalismo Digital”, busca-se uma adaptação ou criação de ordenamentos jurídicos mais territoriais para sanar essas lacunas jurídicas e, também, concretizou-se uma regulamentação internacional, mas de abrangência apenas à União Europeia, que se assemelha ao funcionamento do Comitê, mas possui uma natureza estatal.

Tanto o *Oversight Board* como a DSA possuem um poder de deliberação sobre a aplicação de moderação de conteúdo nas redes sociais, mas enquanto se trata de um órgão estatal, a outra é apenas uma espécie de conselho de uma empresa privada - muito embora, claro, ter assegurada sua prerrogativa de independência.

O cerne aqui é que mesmo o *Oversight Board* ter sido um recurso revolucionário na moderação de conteúdo nas redes sociais, esbarra-se no obstáculo de que as mídias de Zuckerberg possuem a faculdade ou não de acatar o que é decidido nos julgamentos do Comitê. Logo, nada impede de que o “tribunal” seja ignorado ou desfeito caso seja conveniente aos interesses da empresa.

Dessa forma, pode-se concluir que embora o Comitê de Supervisão, atualmente, seja uma ferramenta bastante útil e íntegra para a garantia da proteção do direito humano à liberdade de expressão e tenha realizado uma excelente atuação no julgamento do caso de Donald Trump, o *Board* não possui segurança jurídica alguma, estando à mercê de seu criador para acatar ou não suas deliberações e sugestões, bem como está limitada apenas a um grupo de plataformas digitais ao comando do *Meta Inc.*

Todavia, enquanto o Comitê prossegue em funcionamento, é importante destacar nessa conclusão que a eleição das regulamentações do direito internacional voltados aos direitos humanos, dá ao “órgão julgador privado” uma aparência de maior neutralidade em suas decisões. A aplicação dos princípios internacionais de legalidade, legitimidade, necessidade e proporcionalidade na limitação à liberdade de expressão, auxiliam os julgadores a verificarem no caso concreto o tipo de medida que mais se adequa na moderação de conteúdo, o que seria aplicado em qualquer Corte internacional de direitos humanos, caso fossem designadas a este papel.

Aqui, ressalta-se a consideração de que não se deve ocorrer em nenhum momento, seja por parte de qualquer mídia social ou pelo Comitê de Supervisão, a censura prévia nas manifestações de seus usuários. Mas sendo livres para se expressar dentro do espaço virtual, deve-se ter o discernimento que alguns discursos podem violar o direito de outrem e que, a partir disso, poderão sofrer sanções relativas ao conteúdo postado. Fica-se claro que o usuário não será proibido, em regra, de se manifestar em seu perfil, mas que o conteúdo inapropriado que viole padrões da comunidade, bem como *standarts* de direitos humanos, poderão ser removidos das plataformas.

Volta-se, então, para a averiguação dos princípios limitadores da livre

manifestação que no caso de Trump tiveram que mensurar não somente a violação que seu conteúdo havia incorrido, mas à proporção que seu discurso tomou para além dos limites do ciberespaço, os danos presente e futuros (médio e longo prazo) que ocorreriam se o conteúdo e seu interlocutor se mantivessem ativos na rede social. O que não se trata de medida de censura, mas sim na contenção de prejuízos à segurança da população e, também, da democracia norte-americana.

Evidente que foi desnecessária a suspensão por tempo indeterminado das contas de Trump pelo *Facebook* e *Instagram*, o que foi corrigido pelo Comitê, que recomendou que nenhum usuário pode ter suas contas suspensas *ad aeternum*, devendo ser considerado o tempo de suspensão em conformidade com a necessidade de proporcionalidade que a medida poderá colaborar no controle dos danos.

Portanto, para se ter um efetivo instrumento de garantia à proteção do *free speech* é necessária uma atuação em conjunto do poder público internacional com o setor privado voltado às *Bigtechs*, para que assim possa existir uma maior abrangência de redes sociais e uma melhor fiscalização da moderação de conteúdo.

Como o *Oversight Board* pode garantir a proteção do direito humano à liberdade de expressão, após a análise do cotejo dos Padrões da Comunidade do *Facebook* e os *standards* Internacionais de direitos humanos sobre liberdade de expressão em comparação ao regramento norte-americano, tendo em conta a análise do caso referente à suspensão das contas do *Facebook* e *Instagram* do ex-presidente dos Estados Unidos, Donald Trump? Bom, o Comitê pode não ser a solução do futuro para esse grande problema da Era Digital, mas, sem sombra de dúvidas, é um caminho que será seguido, remodelado e replicado para melhor resguardar o direito fundamental à liberdade de expressão em detrimento à moderação de conteúdo nas mídias sociais.

REFERÊNCIAS

ALVES, Rodrigo Vitorino Souza; RIBEIRO, Gabriel de Almeida ; CARDOSO, Guilherme Gonçalves . Intolerância, Liberdade de Expressão e a incidência do Hate Speech no Mundo Moderno. **Diritto & Diritti**, v. 01, p. 01, 2012.

ARCHEGAS, João Victor et al. **Proteção de Dados e Transparência em Moderação de Conteúdo na Europa, Reino Unido e Brasil**. 2021.

BAKER, Edwin c., Harm, Liberty, and Free Speech, **Southern California Law Review** 70 (1997).

BARLOW, John Perry. **Declaração de Independência do Ciberespaço**. Disponível em: <http://www.dhnet.org.br/ciber/textos/barlow.htm>. Acesso em: 30 maio de 2023.

BARROSO, Luiz Roberto. Neoconstitucionalismo e Constitucionalização do Direito. **Revista de Direito Administrativo**. v. 240, 2005.

BRUGGER, Winfried. Proibição ou proteção do discurso do ódio? Algumas observações sobre o direito alemão e o americano. Tradução de Maria Ângela Jardim de Santa Cruz Oliveira. **Revista de Direito Público**. Brasília: Instituto Brasiliense de Direito Público, ano 4, v. 15, n. 117, jan./mar. 2007.

Capitol riots timeline: what happened on 6 january 2021? BBC News, 2022. Disponível em: <https://www.bbc.com/news/world-us-canada-56004916>. Acesso em 25 jun. 2023.

CASTELLS, Manuel. **A era da informação: economia, sociedade e cultura**. Tradução de Roneide Venancio Majer. São Paulo: Paz e Terra, 2002, v. 1.

COMISSÃO EUROPEIA. **Proposal for a Regulation of the European Parliament and of the Council on a Single Market for Digital Services (Digital Services Act) and amending Directive 2000/31/EC**. Bruxelas, 2020, p. 2. Disponível em: <https://bit.ly/3gWQqvj>. Acesso em: 20 maio.2023.

DINIZ, Célia Regina; SILVA, Iolanda Barbosa da. **Metodologia científica**. Campina Grande: EDUEP, 2008. Disponível em: http://ead.uepb.edu.br/ava/arquivos/cursos/geografia/metodologia_cientifica/Met_Cie_A04_M_WEB_310708.pdf. Acesso em: 26 out. 2021.

DWORKIN, Ronald. **Free Speech and its Limits**. 1992.

EMPOLI, Giuliano de. **Os engenheiros do caos**. 1. ed. São Paulo: Vestígio, 2019. E-book

ESTADOS UNIDOS. **The Constitution of the United States**. Washington, DC, 1791. Disponível em: <https://constitution.congress.gov/constitution/>. Acesso em: 26 out. 2021.

ESTARQUE, Marina; ARCHEGAS, João Victor. **Redes Sociais e Moderação de Conteúdo: Criando regras para o debate público a partir da esfera privada**. Instituto de Tecnologia e Sociedade do Rio de Janeiro, 2020. Disponível em: <https://itsrio.org/wp-content/uploads/2021/07/Protecao-de-Dados-e-Transparencia-em-Moderacao-de-Conte%C3%BAdo.pdf>. Acesso em: 23 maio 2023.

FACHIN, Zulmar. Desafios da regulação do ciberespaço e a proteção dos direitos da personalidade. **Revista Jurídica (FURB)**, [S.l.], v. 25, n. 56, p. e10081, jan. 1970. ISSN 1982-4858. Disponível em: <https://bu.furb.br/ojs/index.php/juridica/article/view/10081>. Acesso em: 30 maio 2023.

FLORIDI, Luciano. **Humanismo em bits**. Entrevista com Luciano Floridi Tradução de Moisés Sbardelotto. INSTITUTO HUMANITAS UNISINOS, 2020. Disponível em: <http://www.ihu.unisinos.br/78-noticias/601785-humanismo-em-bits-entrevista-com-luciano-floridi>. Acesso em: 26 out. 2021.

FLORIDI, Luciano. **The Fourth Revolution: how the infosphere is reshaping human reality**. Oxford University Press, 2014. E-book.

FONSECA, João José Saraiva. **Metodologia da pesquisa científica**. Fortaleza: UEC, 2002.

GREGORIO, Giovanni De. The rise of digital constitutionalism in the European Union. **International Journal of Constitutional Law**, v. 19, ed. 1, Jan. 2021, p. 41–70. Disponível em: <https://doi.org/10.1093/icon/moab001>. Acesso em: 26 out. 2021.

HOWARD, Jeffrey W. Free Speech and Hate Speech. **Annual Review of Political Science**, v. 22, maio/2019. Disponível em: <https://www.annualreviews.org/doi/full/10.1146/annurev-polisci-051517-012343>.

KAPLAN, Andreas. **Social Media, Definition, and History**. Encyclopedia of Social Network Analysis and Mining. LLC, 2016.

KLONICK, Kate. The Facebook Oversight Board: Creating an Independent Institution to Adjudicate Online Free Expression. **Yale Law Journal**, v. 129, ed. 2418, 2020. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3639234. Acesso em: 15 jul. 2023.

LÉVY, Pierre. **Cibercultura**. Tradução de Carlos Irineu da Costa. 3 ed. São Paulo: 34, 2009.

LYND, Staughton. Brandenburg v. Ohio: A Speech Test for All Seasons?. **University of Chicago Law Review**, Vol. 43 : Iss. 1 , Article 34, 1975. Disponível em: <https://chicagounbound.uchicago.edu/uclrev/vol43/iss1/34>. Acesso em: 26 out. 2021.

MARCONI, Marina de Andrade; LAKATOS, Eva Maria. **Fundamentos da Metodologia Científica**. 5ª ed. São Paulo: Atlas, 2003. Disponível em: https://docente.ifrn.edu.br/olivianeta/disciplinas/copy_of_historia-i/historia-ii/china-e-india. Acesso em: 26 out. 2021.

MEYER-PFLUG, Samantha Ribeiro. **Liberdade de Expressão e Discurso do Ódio**. São Paulo: Editora Revista dos Tribunais, 2009.

MINOW, Martha. **Regulating Hatred Whose Speech, Whose Crimes, Whose Power**. 2000.

MODELLI, Laís. **Relembre as mentiras mais famosas de Trump**. G1, 2020. Disponível em: <https://www.cnbc.com/2021/01/13/trump-tweets-legacy-of-lies-misinformation-distrust.html>. Acesso em: 25 jun. 2023

NEGÓCIO, Ramon de Vasconcelos. Autorregulação digital: Da normatividade excludente para o diálogo normativo com o Estado. **Revista Culturas Jurídicas**, Vol. 10, mar. 2022. Disponível em: <https://periodicos.uff.br/culturasjuridicas/article/view/48188>. Acesso em: 30 maio 2023.

OCCIUZZI, Rita de Cássia Moreno de Souza. **O direito à liberdade de expressão e o discurso de ódio a democracia contemporânea**. 2017. 102f. Dissertação (Mestrado em Direito Constitucional) – Universidade de Fortaleza, Fortaleza, 2017.

ONU. Organização das Nações Unidas. **Declaração Universal dos Direitos Humanos**, 1948. Disponível em: <https://www.unicef.org/brazil/declaracao-universal-dos-direitos-humanos>. Acesso em: 26 out. 2021.

ONU. Organização das Nações Unidas. **Pacto Internacional sobre os Direitos Civis e Políticos**, 1966. Disponível em: <https://www.oas.org/dil/port/1966%20Pacto%20Internacional%20sobre%20Direitos%20Civis%20e%20Pol%C3%ADticos.pdf>. Acesso em: 26 out. 2021.

ONU. Assembleia Geral da Organização das Nações Unidas. **Promotion and protection of the right to freedom of opinion and expression**. 9 out. 2019.

OEA. Organização Dos Estados Americanos. **Convenção Americana sobre Direitos Humanos**, 1969. Disponível em: https://www.cidh.oas.org/basicos/portugues/c.convencao_americana.htm. Acesso em: 26 out. 2021.

OEA. Organização Dos Estados Americanos. **Marco Jurídico Interamericano del derecho a la libertad de expresión**. Comissão Interamericana de Direitos Humanos, 2010. Disponível em: <http://www.oas.org/es/cidh/expresion/docs/publicaciones/MARCO%20JURIDICO%20INTERAMERICANO%20DEL%20DERECHO%20A%20LA%20LIBERTAD%20DE%20EXPRESION%20ESP%20FINAL%20portada.doc.pdf>. Acesso em: 26 out. 2021.

RATTNER, Nate. **Trump's election lies were among his most popular tweets.** CNBC, 2021. Disponível em: <https://www.cnn.com/2021/01/13/trump-tweets-legacy-of-lies-misinformation-distrust.html>. Acesso em: 25 jun. 2023.

RECUERO, Raquel. **Redes sociais na internet.** Porto Alegre: Sulina, 2009. *E-book*

RIBEIRO, Raisal Duarte da Silva; COSTA, Rodrigo de Souza. Discurso de ódio: uma análise comparada dos ordenamentos jurídicos alemão, brasileiro e estadunidense. **Direitos e garantias fundamentais II** [Recurso eletrônico on-line], organização CONPEDI/UNICURITIBA; Coordenadores: Marcos Leite Garcia, Matheus Felipe De Castro – Florianópolis: CONPEDI, 2016.

SCHWAB, Klaus. **A Quarta Revolução Industrial.** Tradução de Daniel Moreira Miranda. São Paulo: Edipro, 2016. *E-book*

SILVA, Isabel Germano Rodrigues; SILVA, Josiane da Costa. Liberdade de expressão e seus limites: discurso de ódio é tolerável? **VirtuaJus**, Belo Horizonte, v. 3, n. 5, p. 255-273, 2º sem. 2018. Disponível em: <http://periodicos.pucminas.br/index.php/virtuajus/article/view/19519>. Acesso em: 26 out. 2021.

SILVEIRA, Denise Tolfo; CÓRDOVA, Fernanda Peixoto. **Métodos de Pesquisa.** Organizado por Tatiana Engel Gerhardt e Denise Tolfo Silveira. Porto Alegre: Editora UFRGS, 2009. Disponível em: <http://www.ufrgs.br/cursopgdr/downloadsSerie/derad005.pdf>. Acesso em: 26 out. 2021.

STONE, Geoffrey R. Content-neutral restriction. *University Chicago Law Review*, n. 54, 1987, p. 46-118. Disponível em: https://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=2968&context=journal_articles. Acesso em: 26 out. 2021.

STONE, Geoffrey R. Hate Speech and the U.S. Constitution. **East European Constitution Review**. n. 3, 1994, p. 78-82. Disponível em: https://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=3246&context=journal_articles. Acesso em: 26 out. 2021.

SUZOR Nicolas. Digital Constitutionalism: Using the Rule of Law to Evaluate the Legitimacy of Governance by Platforms. **Social Media + Society**. July 2018. Disponível em: <https://journals.sagepub.com/doi/full/10.1177/2056305118787812>. Acesso em: 26 out. 2021.

TAKANO, Camila Cardoso; SILVA, Lucas Gonçalves da. O constitucionalismo digital e as novas tecnologias da informação e comunicação (TIC). **Revista de Direito, Governança e Novas Tecnologias**. v. 6. n. 1. jan/jun. 2020. Disponível em: <https://indexlaw.org/index.php/revistadgnt/article/view/6392/pdf>. Acesso em: 26 out. 2021.

TELLES, André. **A revolução das mídias sociais: cases, conceitos, dicas e ferramentas.** São Paulo: M. Books, 2010.

TORTELLA, Tiago. **Invasão do Capitólio completa um ano: relembre o ataque à democracia dos EUA**. CNN Brasil, 2022. Disponível em: <https://www.cnnbrasil.com.br/internacional/invasao-ao-capitolio-completa-um-ano-relembre-o-ataque-a-democracia-dos-eua/>. Acesso em: 25 jun. 2023.

TRACKSTON, Mark C. WISCONSIN v. MITCHELL 113 S.Ct. 2194 (1993) United States Supreme Court, 1 **Race & Ethnic Anc.** L. Dig. 65 (1995). Disponível em: <https://scholarlycommons.law.wlu.edu/crsj/vol1/iss1/16>. Acesso em: 07.05.2023

WALDRON, Jeremy. **The Harm in Hate Speech**. Cambridge (Mass.): Harvard University Press, 2012.

WE ARE SOCIAL. Hootsuite Inc. **Digital 2022**: Global Overview Report. 2022. Disponível em: <https://wearesocial.com/uk/blog/2022/01/digital-2022-another-year-of-bumper-growth-2/>. Acesso em: 03 mar. 2022.

WE ARE SOCIAL. Meltwater Inc. **Digital 2023**: Global Overview Report. 2023. Disponível em: <https://wearesocial.com/uk/blog/2023/01/the-changing-world-of-digital-in-2023/>. Acesso em: 01 ago. 2023.

WERTHEIN, Jorge. A sociedade da informação e seus desafios. **Ciência da Informação**, Brasília, v. 29, n. 2, p. 71-77, maio/ago. 2000.

WOLKMER, Antônio Carlos. DIREITOS HUMANOS: Novas Dimensões e Novas Fundamentações. **Revista Direito Em Debate**, n. 16/17, jan./jun., 2002. Disponível em: <https://www.revistas.unijui.edu.br/index.php/revistadireitoemdebate/article/view/768>. Acesso em: 26 out. 2021.

ZIBETTI, Fabiola Wust; PILAU SOBRINHO, Liton Lanes; MIRANDA, Luiz Ricardo de. A influência do critério da soberania nas fontes do Direito Internacional. **Revista Jurídica**, Curitiba, v. 03, n. 52, 2018, p. 106-128.