

III ENCONTRO VIRTUAL DO CONPEDI

DIREITO, GOVERNANÇA E NOVAS TECNOLOGIAS II

JOSÉ RENATO GAZIERO CELLA

AIRES JOSE ROVER

AGATHA GONÇALVES SANTANA

Todos os direitos reservados e protegidos. Nenhuma parte deste anal poderá ser reproduzida ou transmitida sejam quais forem os meios empregados sem prévia autorização dos editores.

Diretoria - CONPEDI

Presidente - Prof. Dr. Orides Mezzaroba - UFSC - Santa Catarina

Diretora Executiva - Profa. Dra. Samyra Haydêe Dal Farra Napolini - UNIVEM/FMU - São Paulo

Vice-presidente Norte - Prof. Dr. Jean Carlos Dias - Cesupa - Pará

Vice-presidente Centro-Oeste - Prof. Dr. José Querino Tavares Neto - UFG - Goiás

Vice-presidente Sul - Prof. Dr. Leonel Severo Rocha - Unisinos - Rio Grande do Sul

Vice-presidente Sudeste - Profa. Dra. Rosângela Lunardelli Cavallazzi - UFRJ/PUCRio - Rio de Janeiro

Vice-presidente Nordeste - Profa. Dra. Gina Vidal Marcilio Pompeu - UNIFOR - Ceará

Representante Discente:

Prof. Dra. Sinara Lacerda Andrade - UNIMAR/FEPODI - São Paulo

Conselho Fiscal:

Prof. Dr. Caio Augusto Souza Lara - ESDHC - Minas Gerais

Prof. Dr. João Marcelo de Lima Assafim - UCAM - Rio de Janeiro

Prof. Dr. José Filomeno de Moraes Filho - Ceará

Prof. Dr. Lucas Gonçalves da Silva - UFS - Sergipe

Prof. Dr. Valter Moura do Carmo - UNIMAR - São Paulo

Secretarias

Relações Institucionais:

Prof. Dra. Daniela Marques De Moraes - UNB - Distrito Federal

Prof. Dr. Horácio Wanderlei Rodrigues - UNIVEM - São Paulo

Prof. Dr. Yuri Nathan da Costa Lannes - Mackenzie - São Paulo

Comunicação:

Prof. Dr. Liton Lanes Pilau Sobrinho - UPF/Univali - Rio Grande do Sul

Profa. Dra. Maria Creusa De Araújo Borges - UFPB - Paraíba

Prof. Dr. Matheus Felipe de Castro - UNOESC - Santa Catarina

Relações Internacionais para o Continente Americano:

Prof. Dr. Heron José de Santana Gordilho - UFBA - Bahia

Prof. Dr. Jerônimo Siqueira Tybusch - UFSM - Rio Grande do Sul

Prof. Dr. Paulo Roberto Barbosa Ramos - UFMA - Maranhão

Relações Internacionais para os demais Continentes:

Prof. Dr. José Barroso Filho - ENAJUM

Prof. Dr. Rubens Beçak - USP - São Paulo

Profa. Dra. Viviane Coêlho de Séllos Knoerr - Unicuritiba - Paraná

Eventos:

Prof. Dr. Antônio Carlos Diniz Murta - Fumec - Minas Gerais

Profa. Dra. Cinthia Obladen de Almendra Freitas - PUC - Paraná

Profa. Dra. Livia Gaigher Bosio Campello - UFMS - Mato Grosso do Sul

Membro Nato - Presidência anterior Prof. Dr. Raymundo Juliano Feitosa - UMICAP - Pernambuco

D597

Direito, governança e novas tecnologias II [Recurso eletrônico on-line] organização CONPEDI

Coordenadores: Agatha Gonçalves Santana; Aires Jose Rover; José Renato Gaziero Cella – Florianópolis: CONPEDI, 2021.

Inclui bibliografia

ISBN: 978-65-5648-320-7

Modo de acesso: www.conpedi.org.br em publicações

Tema: Saúde: segurança humana para a democracia

1. Direito – Estudo e ensino (Pós-graduação) – Encontros Nacionais. 2. Governança. 3. Novas tecnologias. III Encontro Virtual do CONPEDI (1: 2021 : Florianópolis, Brasil).

CDU: 34



III ENCONTRO VIRTUAL DO CONPEDI

DIREITO, GOVERNANÇA E NOVAS TECNOLOGIAS II

Apresentação

DIREITO, GOVERNANÇA E NOVAS TECNOLOGIAS II

O III Encontro Virtual do Conselho Nacional de Pesquisa e Pós-Graduação em Direito (CONPEDI), realizado ainda no decorrer da pandemia do COVID-19, aponta para temas que indicam o início de uma grande revolução tecnológica que atinge o mundo todo em seus aspectos político, jurídico, social e econômico. A quarta revolução industrial, citada ao longo da grande maioria dos trabalhos, mais do que demonstrar a necessidade de um olhar sob uma óptica transdisciplinar, aponta para a necessidade de repensar muitas das bases científicas que se estudou anteriormente, revisitando muitos dos institutos tradicionais do direito para o cumprimento de seu mister de realização do bem comum.

Dentre os temas destacados nos trabalhos, pode-se perceber as mudanças no modo de existir das relações humanas, da comunicação, a preocupação com os dados pessoais disponibilizados e gerenciados não apenas pelos provedores de internet como pelo próprio Estado, tecnologias disruptivas, todas trabalhadas dentro do contexto do atual fenômeno da tecnoglobalização, buscando soluções para problemas que se apresentam e para aqueles que, assim como distopias descritas na literatura do passado, parecem se desenhar em um futuro muito próximo.

Por conta dessa pluralidade de visões e problemáticas trazidas pelos autores, os artigos apresentados foram agregados em três blocos, com o escopo de aprofundar o debate sobre temas tão caros e complexos, criando assim um fio condutor para o grupo de trabalho, em um desdobramento lógico.

No primeiro bloco dos trabalhos, os temas centraram-se no debate sobre a governança, essencialmente relacionada à gestão dos dados pessoais e as reflexões sobre a aplicabilidade da lei de proteção de dados dentro do espaço virtual, temas hoje muito caros, essencialmente frente ao atual estado da arte da tecnologia mundial advindo com a pandemia. Nesse sentido, foram abordadas em análises principiológica e legislativa, a partir de abordagens teóricas e empíricas sobre as problemáticas da vigilância governamental; governança sobre bancos de dados de crédito; riscos relacionados ao uso de dados pessoais dentro da prática da telemedicina; e os impactos dos algoritmos criados pelas grandes empresas da rede mundial de computadores.

Logo em seguida, no segundo bloco, o núcleo dos artigos gravita em torno das novas tecnologias emergentes aplicadas tanto dentro dos ambientes e instituições públicos quanto privados, tendo destaque reflexões críticas sobre a tecnologia blockchain como meio de conferir maior segurança e imutabilidade de dados; reconhecimento de dados biométrico; nanotecnologia; processos decisórios automatizados e transparência algorítmica. Todos os artigos trazem à baila a necessidade de aprofundamento e diálogo com outras áreas de conhecimento para um redesign de muitas das estruturas sociais e sociedades em rede hoje conhecidas.

O terceiro e derradeiro bloco foi dividido essencialmente tendo em vista temas multidisciplinares correlatos à justiça dentro do contexto do direito, governança e novas tecnologias, destacando a necessidade de aprimoramento e proteção sobre as inovações, que devem ser vistas como forma de garantia de efetivação de direitos e combate às ilicitudes e a promoção da prevenção e reparação de danos. Assim, são abordados temas sensíveis como fake news e discurso de ódio nas redes; big techs; uma visão comparada do direito ao esquecimento no Brasil e na Europa; pornografia de vingança; transparência fiscal na responsabilidade civil e a corrupção sob o aspecto da governança e reflexões sobre a herança digital no Brasil.

Todos os artigos configuram estudos de excelência na área, e seu compartilhamento representa grande contribuição e referência para estudantes, pesquisadores e demais profissionais do direito e de outras áreas de conhecimento. Assim, os coordenadores desse grande grupo de trabalho convidam a todos a ler na íntegra os artigos no sentido de fomentar e ampliar o diálogo, o debate e as pesquisas nessas temáticas que compõem problemas atuais e possíveis em um futuro próximo, dentro da realidade do mundo contemporâneo.

Prof. Dr. Aires José Rover - UFSC

Prof. Dr. José Renato Gaziela Cella - IMED

Prof.a Dra. Agatha Gonçalves Santana - UNAMA

O DIREITO HUMANO À LIBERDADE DE EXPRESÃO E AS INICIATIVAS DE CONTROLE DAS FAKE NEWS E DO DISCURSO DE ÓDIO PELAS REDES SOCIAIS

THE HUMAN RIGHT TO FREEDOM OF EXPRESSION AND CONTROL INITIATIVES FOR FAKE NEWS AND HATE SPEECH BY SOCIAL MEDIAS

**Fernanda da Luz Berna
Gustavo Silveira Borges**

Resumo

O artigo tem por objetivo geral estudar o direito humano à liberdade de expressão e as iniciativas de controle das fake news e do discurso de ódio pelas redes sociais. Buscou-se traçar um panorama entre as consequências da hiperconexão na era digital com as iniciativas para conter a disseminação de fake news e do discurso de ódio nas redes sociais. Optou-se pela pesquisa qualitativa e pelo método dedutivo. Conclui-se que apesar de a liberdade de expressão ser um direito fundamental, esta possui limites, os quais devem ser respeitados no momento de manter ou retirar conteúdo impróprio nas plataformas digitais.

Palavras-chave: Liberdade de expressão, Direitos humanos, Fake news, Discurso de ódio, Redes sociais

Abstract/Resumen/Résumé

This article aims to study the human right to freedom of expression and the initiatives to control fake news and hate speech through social networks. We sought to draw a picture of the consequences of hyperconnection in the digital age with initiatives to contain the spread of fake news and hate speech on social medias. Qualitative research and the deductive method were chosen. It is concluded that although freedom of expression is a fundamental right, it has limits, which must be respected when maintaining or removing inappropriate content on digital platforms.

Keywords/Palabras-claves/Mots-clés: Freedom of liberty, Human rights, Fake news, Hate speech, Social medias

INTRODUÇÃO

O presente artigo irá abordar um estudo acerca do direito humano à liberdade de expressão e as iniciativas de controle das *fake News* e do discurso de ódio pelas redes sociais, contextualizando com essas novas tecnologias advindas sobretudo a partir da expansão da internet como a inteligência artificial, a robótica, realidade virtual, neurotecnologias, entre outros, as quais constituem a Quarta Revolução Industrial (SCHWAB; DAVIS, 2019, p. 32).

É inevitável que o período histórico em que a sociedade pós-moderna está vivenciando vem acompanhada de diversos benefícios e malefícios: a possibilidade de expansão da liberdade dentro do ambiente virtual, a ampliação do ambiente democrático, melhores níveis de saúde e educação, são algumas das possibilidades (SCHWAB; DAVIS, 2019, p. 38). Quanto aos impactos negativos, Schwab e Davis (2019, p. 63) compartilham diversas preocupações atreladas às tecnologias, como fragmentação política, incertezas econômicas, desinformação e outras infinitas possibilidades; o terreno de potenciais utilizações é demasiadamente fértil diante da versatilidade tecnológica, acarretando riscos por vezes imprevisíveis. Outra problemática se insere no âmbito eleitoral; os paradigmas das campanhas eleitorais modificaram-se para fragmentar os cidadãos através de campanhas direcionadas baseadas em seus dados pessoais, e assim instituiu-se uma nova política baseada na ideia de que “para conquistar uma maioria, eles não vão convergir para o centro, e sim unir-se aos extremos” através de propagandas que façam emergir sobretudo emoções negativas (EMPOLI, 2019, p. 13).

Nesse interim, para enfrentar os desafios advindos com as novas tecnologias é de extrema importância resgatar valores sociais universais como a dignidade humana, o reconhecimento de bens comuns em detrimento de valores individualistas e a necessidade de mecanismos de gerenciamento do ambiente virtual (SCHWAB; DAVIS, 2019, p. 62). A complexidade dos novos fenômenos e seus impactos na sociedade devem ser abordados de forma multilateral e enseja a participação de diversos atores; é necessário a colaboração da área acadêmica, do governo, de líderes empresariais, sociedade civil e organizações internacionais (SCHWAB; DAVIS, 2019, p. 78). Diante disso, o presente trabalho busca analisar os mecanismos de controle que estão sendo implementados pelas mídias sociais do *Facebook* e do *Twitter* para restringir a disseminação de *fake news* e do discurso de ódio postados por seus usuários. Busca-se, também, verificar se, nessas

ferramentas, a remoção dos conteúdos não violam o direito humano e fundamental à liberdade de expressão.

Com efeito, o problema de pesquisa que norteou a presente investigação é o seguinte: considerando o impacto negativo que as *fake news* e o discurso de ódio propagado dentro das redes sociais estão causando na sociedade pós-moderna, pode-se dizer que a tomada de medidas de controle a disseminação desse conteúdo pelas redes sociais estaria violando o direito humano à liberdade de expressão?

De início, é imprescindível compreender que as *fake news* sempre estiveram presentes em nossa sociedade, bem como o discurso de ódio, mas com o avanço tecnológico, as mídias sociais e a hiperconexão que o mundo está inserido atualmente, faz com que o alcance e a rapidez da reprodução do conteúdo postado sejam altamente perigosos, uma vez que mobilizam milhares de pessoas de determinado grupo social a se rebelar e, inclusive, atentar contra minorias e, também, contra o estado democrático de direito.

Assim, tornou-se de extrema importância a busca, não somente por legislações para controlar esse tipo de conteúdo nocivo nas redes sociais, mas por uma iniciativa própria das *social medias* que possa filtrar e remover os conteúdos falsos e de discurso de ódio, procurando compreender a aplicação desse direito à liberdade de expressão, o qual vem sendo defendido nas legislações internacionais desde o término da Segunda Guerra Mundial e o fim do regime nazista, e os seus limites legais.

Trata-se de pesquisa qualitativa, que se utilizou de dados documentais e doutrinários referentes ao tema da pesquisa. Adotou-se também os casos *The Oversight Board* e *Birdwatch* como modelo de análise às iniciativas de controle à disseminação de *fake news* e do discurso de ódio nas redes sociais. O método de pesquisa utilizado foi o hipotético-dedutivo.

A narrativa foi desenvolvida da seguinte forma: na primeira etapa, a pesquisa analisa a sociedade da informação, buscando identificar a inserção das *fake news* dentro dela. Na sequência, volta-se ao estudo do emprego do discurso de ódio nas redes sociais, bem como a ilustração de algumas consequências atuais ocorridas nos últimos tempos. E, por fim, examina a aplicação do direito à liberdade de expressão e seus limites legais na criação de mecanismos para controlar o uso e disseminação das notícias falsas e no discurso de ódio nas redes sociais, tendo como exemplo os casos *The Oversight Board*, do *Facebook*, e *Birdwatch*, do *Twitter*.

1. A ERA DIGITAL E AS *FAKE NEWS*

A civilização humana passou por diversas mudanças e inovações como sociedade, desde a revolução agrícola, há mais de dez mil anos, na qual se pode permitir a produção de alimentos, crescimento populacional e a formação de pequenas comunidades (SCHWAB, 2016, p. 18), até os tempos atuais. Nossa sociedade passou por três revoluções industriais, as quais foram grande caminho para estar inserida hoje em um contexto altamente avançado em termos de tecnologia, produção e ciência, dentre outras áreas. (SCHWAB, 2016, p. 18-19)

Entretanto, quando se pensava que havia uma estagnação após a 3ª revolução, a qual é conhecida como revolução digital ou do computador, haja vista ser sido fomentada computação e pela internet, Schwab (2016, p. 18-19) acredita que estamos no início da quarta revolução industrial, onde pode ser caracterizada “por uma internet mais ubíqua e móvel, por sensores menores e mais poderosos que se tornam mais baratos e pela inteligência artificial e aprendizagem automática (ou aprendizado de máquinas)”. Segundo, ainda Schwab (2016, p. 19-20), essa nova revolução industrial que estamos vivendo vai muito além de sistemas, máquinas inteligentes e conectadas; ela se trata da difusão da tecnologia, nas mais diversas áreas, de uma forma mais célere e de longo alcance às mais distintas partes do mundo.

Toda rapidez e dimensão que esse progresso tecnológico acontece é amplamente visível por todos nós, basta observar as tecnologias aplicadas nos smartphones, aplicativos, sejam eles de pesquisa, redes sociais e financeiros; na ciência, com bioética e, atualmente, a busca pela vacina na Covid-19; bem como na aplicação da tecnologia para um desenvolvimento mais sustentável na sociedade, seja através de energias renováveis até no campo da moda com a chamada *ecofashion*, inclusive.

Todavia, mesmo diante de tantos avanços em diferentes áreas, nessa *era digital* o que mais nos atinge diretamente e influencia em nosso dia a dia são as redes sociais, as quais são tidas como “facilitadores de interação social com pessoas que, a princípio possuem relações próximas, seguidas de aproximação por interesse nos mesmos conteúdos” (MENDONÇA, 2019, p. 47), esse contexto é muito bem observado, também, por Schwab (2016, p. 100) que complementa o entendimento de que as mídias digitais estão cada vez mais orientando o posicionamento do cidadão, seja ele de forma individual ou coletiva, diante da sociedade, justamente porque as conexões interpessoais que aconteciam de uma pessoas para uma outra diretamente, agora passa a ser de uma pessoa

para milhares de outras, o que permite a construção de novos laços e a sua manutenção entre aqueles grupos que possuem a mesma opinião, independentemente da distância de onde estiveram, ultrapassando “os limites, sociais, econômicos, culturais, políticos religiosos e ideológicos.

Sim, a sociedade da *era digital* não para, ela está conectada o tempo todo com todo o mundo, tendo sido nomeada por Jonathan Crary de sociedade *non-stop*, se caracterizando pela “expressão 24/7 – vinte e quatro horas por dia – sete dias da semana ou *sempre ligado* – a qual serve como índice de uma modernidade cada vez mais hiperconectada, acelerada e atravessada pelas redes de produção, financeirização e comunicação ininterruptas, e também pela dependência de máquinas eletrônicas”. (WILKE, 2020, p. 10).

Contudo, apesar dos imensos benefícios que a hiperconexão sem pausa trouxe ao mundo, ela acabou reavendo e aprimorando, também, uma face prejudicial, que nos últimos anos se intensificou cada vez mais nas redes sociais: as *fake news* e a disseminação do discurso de ódio. É claro que esses ônus não foram criados a partir da *quarta revolução industrial*, é uma prática antiga e que pode ser muito estudada e observada, por exemplo, durante a *segunda guerra mundial* com o regime nazista, liderado por Adolf Hitler. No entanto, o que se demorava por muito mais tempo o alastramento de notícias falsas e a pregação de um discurso de ódio, nos dias atuais está acessível para qualquer pessoa na tela de seu celular e a qualquer tempo.

Os desenvolvedores das redes sociais (*Facebook, Twitter, Instagram* e outros) criaram, ao longo dos anos, mecanismos para que cada vez mais nossa atenção esteja voltada a esses aplicativos, haja vista que tempo gasto pelos usuários lhes dá o que mais almejam: o lucro. Ocorre que no meio desse “caminho” entre o nosso engajamento até a incorporação de capital, os *softwares* colhem os dados cada usuário. O questionamento que fica é para onde vão esses dados e por quem são usados? (O DILEMA DAS REDES, 2020)

Uma amostra de quanto essa coleta de dados pode ser perigosa é, manifestamente, verificada no escândalo do vazamento de dados do *Facebook* pela *Cambridge Analytica*, em que a empresa de consultoria utilizou de um aplicativo de enquete, o *thisisyourdigitallife*, para colher não somente os dados daqueles usuários que aceitavam os termos e condições sem ler, mas, também, de seus amigos na rede social. De acordo com o noticiado, pelo menos 270 mil pessoas fizeram o teste, no entanto, foram coletados os dados um total de 50 milhões de usuários, em sua maioria norte-americanos

(BBC NEWS, 2018), com o intuito de traçar perfis de personalidade dos usuários, criando, assim, nichos de acordo com a qualidade encontrada no *quiz* e direcionando a eles a exibição de anúncios com fins políticos que provocassem certa simpatia ou conexão com o determinado assunto (DEMARTINI, 2018).

Toda essa “engrenagem” das redes sociais acaba se tornando uma manipulação em bolhas e que muitas vezes levam seus usuários para a polarização política. O documentário *O Dilema das redes* (2020), traz, também, o exemplo do *youtube*, onde o usuário começa escolhendo vídeos sobre terraplanismo e não se dá conta que, a partir daí, começará a ser recomendado uma série de vídeos de teorias conspiratórias, inclusive, outros tipos de notícias falsas. Segundo os próprios desenvolvedores dos sistemas das *social medias*, acabou-se criando um sistema inclinado às informações falsas, não intencionalmente, mas se constatou que essas informações inverídicas acabavam levando mais lucro para as empresas do que as verdadeiras, ou seja, a desinformação passou a gerar lucro. (O DILEMA DAS REDES, 2020.)

As mídias sociais possuem uma função pós-massiva, cuja característica é ressaltada por permitir “a personalização e propagação da informação sem o controle de empresários, Estado ou marketing” (LEMOS, 2007, p. 91 apud SANTOS; PAES; PONTES, 2019, p. 154-155), isto é, essa modalidade moderna de mídia, a qual rompe com o fluxo tradicional de informações dos veículos de massa, permite aos receptores “a liberdade de exercer a função de produtores, emissores e coatores de conteúdo”. (GONÇALVES; LOUREIRO, 2018, p. 22 apud SANTOS; PAES; PONTES, 2019, p. 154). Sendo assim, essa função que tornou o usuário o centro da propagação de informações nas mais diversas formas e plataformas de conteúdo, acaba se tornando um facilitador de interação social com pessoas que, pressupõe-se possuem relações próximas, seguidas de aproximação por interesse nos mesmos conteúdos, através de fóruns de discussão e debates, possuindo também outros conteúdos, inclusive comerciais, obviamente (MENDONÇA, 2019, p. 47).

E nessa conjuntura, é importante destacar, que as *fake news* estão inseridas dentro de fenômenos que são identificáveis e intensificados na era da pós-verdade, no entanto, os dois conceitos, apesar de serem associados com frequência, não podem ser confundidos. Enquanto a palavra pós-verdade, eleita como palavra do ano em 2016 pelo Dicionário Oxford, é conceituada como um adjetivo “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” (MCINTYRE, 2018, p.1),

as *fake news* pode sem ser destacadas, segundo Braga (2018, p. 205), como “a disseminação, por qualquer meio de comunicação, de notícias sabidamente falsas com o intuito de atrair a atenção para desinformar ou obter vantagem política ou econômica”, tendo como finalidade o dolo de desinformar o usuário, através de “manchetes desonestas sensacionalistas ou escritas de forma a aumentar as visualizações, os compartilhamentos online e o ganho por cliques na Internet” (WILKE, 2020, p. 13).

Não obstante a existência do direito à liberdade de informação, isto é, de receber e elaborar informações, independentemente do veículo utilizado, garantida ao usuário – e cidadão – o poder de acessar qualquer conteúdo de notícia, “é essencial que se compreenda que nem tudo o que está na internet é de fato útil e benéfico para todos, além do entretenimento” (MENDONÇA, 2019, p. 65), além do mais, é notório que no ambiente virtual contemporâneo há uma enorme produção de informação, as quais não estão sujeitas a um controle de qualidade, o que resulta em um cenário de poluição informacional através da propagação de *Fake news* (ALVES, 2019, p. 270) e segundo Branco (2017, p. 58), o excesso dessas informações constituem um obstáculo para a verificação de todo o conteúdo do usuário, fragmentando sua atenção. Logo, todo o contexto do ambiente das redes sociais vinculado a problemática das *Fake news* interferem, demasiadamente, na concretização ideal do direito à liberdade de expressão; a falta de um controle de qualidade e a rápida propagação de notícias dificulta a categorização de informações dúbias como *Fake news*, uma vez que existe uma variada gama de conteúdo (ALVES, 2019, p. 274).

Para Oliveira e Gomes (2019, p. 107), a propagação de notícias falsas se configura como abuso ao exercício do direito à liberdade de expressão, ferindo o direito informacional da sociedade e o próprio exercício das liberdades a medida que interfere na formação da opinião pública. As mídias sociais amplificaram tanto as *fake News* até ao ponto em que os usuários não mais conseguem distinguir o que é verdade ou mentira, e não importa qual seja o assunto, sempre haverá dúvida e isso possibilita a disseminação de narrativas manipuladoras com muita facilidade (O DILEMA DAS REDES, 2020).

A propagação das notícias falsas é um tema que vêm sendo tratado com extrema relevância na atualidade, justamente porque se tem a compreensão de que essa interferência na formação de opinião, além de refletir “severamente nas escolhas individuais e coletivas, afeta diretamente a esfera pública e privada das pessoas, e, conseqüentemente, o processo democrático” (MENDONÇA, 2019, p. 84). De fato, não há como vislumbrar a possibilidade de existir uma consonância em todas as ideologias,

vivências e dogmas, é utópico demais e, inclusive, podendo se tornar uma distopia caso houvesse um consenso de ideais extremistas. Para Mendonça (2019, p. 88), as divergências são comuns e desejadas, haja vista que é “através dos embates de modos de compreensão dos problemas emergentes é que se abrem as discussões até que se obtenha, nem sempre um consenso ou acordo (que nem sempre é possível), mas uma esperada e desejada decisão e ações coletivas”.

Entretanto, muito embora o dissenso traga em si um ideal democrático e de garantias individuais e coletivas, não se pode ignorar que na era pós-verdade o uso dessa liberdade, de forma mal-intencionada, somado à disseminação de informações falsas na internet trazem enorme prejuízo à sociedade, pois muitos desses conteúdos se voltam, também, para a propagação de um discurso de ódio difundido nas redes sociais, o qual já causou rupturas democráticas significativas no passado e que podem fazer a história se repetir mais uma vez.

2. O DISCURSO DE ÓDIO NO CONTEXTO DAS REDES SOCIAIS

A menos de duas décadas, pelo menos aqui no Brasil, utilizava-se a biblioteca física para pesquisa, mantinha-se amizades e laços através do envio de cartas pelos correios, para realizar ligações e recebe-las era preciso estar em casa, usava-se cheque para pagamentos pré-datados, para pagar contas precisava-se ir até uma agência bancária. Não há dúvida que essa nova era cibernética, com conexões velozes e acessibilidade virtual transformou a vida de nossa sociedade, hoje temos tudo o que precisamos na tela de um telefone móvel.

No entanto, como já havia se falado anteriormente, toda essa tecnologia e inovação não resultou apenas mudanças positivas, mas também intensificou condutas negativas que estão inseridas no fenômeno da era da pós-verdade, tais quais as *fake news* e o discurso de ódio. Essas condutas não são novidades, elas existem há séculos e são utilizadas das mais diversas formas por aqueles que ocupam posições de poder, com um intuito de conquista e obtenção de lucro.

Não existe dúvidas que as redes sociais se tornaram um facilitador de comunicação da atualidade, mas em contrapartida a disseminação de notícias falsas e os mais variados discursos de ódios em grande alcance mundial estão trazendo diversos prejuízos para a sociedade pós-moderna. Segundo Giordan e Massi (2018, p. 280-281) cenário em que somente a grande imprensa e os políticos estivessem no centro do que

chamam de “falseamento ou distorção da realidade” mudou. Muito embora essas duas figuras tenham destaque no direcionamento desses fenômenos, as pessoas comuns acabaram por sair de seus papéis de receptores e, inclusive de “vítimas e alvos passivos” para “exercerem papel ativo em boa parte dos processos que alimentam e dão forma à pós-verdade”, alterando, de certa forma, aquela concepção de que a primazia da realidade seria um pilar essencial da “sociedade e das decisões de interesse público e privado”.

Realmente o modo como surgem e se expandem o discurso de ódio possuem hoje uma variedade de agentes disseminadores, diferentemente do que acontecia, por exemplo, no nazismo, em que se dependia exclusivamente do governo para a criação e a disseminação das notícias falsas e do discurso antissemita, no qual se utilizavam, claro, o que havia disponível como meios de propagação naquela época, tais quais “de transporte público, murais e em qualquer lugar que reunisse concentração de pessoas” (QUESADA; PISA, 2018, p. 1). É óbvio que não se pode esquecer que para que o discurso de ódio tenha potencial propagador é necessário que esse sentimento já exista em determinado grupo social, para que assim possa se alastrar à medida que as pessoas vão se identificando com o discurso empregado.

Segundo Empoli (2019, p. 46), muito embora o discurso de ódio seja facilmente identificável e acessível através das redes sociais, essa ferramenta não foi desenhada, naturalmente, para semear teorias conspiratórias, no entanto, em algum momento se descobriu que a utilização delas para esse propósito seria muito eficiente, já que ajudam na provocação de emoções fortes, indignações, ódio, polêmicas e polarizações, gerando mais engajamento do que uma notícia verdadeira¹. E para a verdade ser menos consumida do que as notícias falsas ou o discurso de ódio, Braga (2018, p. 97) complementa que o problema está na frustração do ser humano na era pós-contemporânea, na qual causou um “crescente abismo entre a mediocridade de nossa vida e todas as vidas possíveis que se oferecem virtualmente em nossos monitores e telas de celular”, fazendo com que sites conspiratórios tenham o condão de exercer um “poder de fascinação intenso porque oferecem, enfim, uma explicação plausível para as dificuldades nas quais nos encontramos” (BRAGA, 2018, p. 45-46).

¹ Um recente estudo do Instituto de Tecnologia de Massachusetts (MIT) demonstrou que uma falsa informação tem, em média, 70% a mais de probabilidade de ser compartilhada na internet, pois ela é, geralmente, mais original que uma notícia verdadeira. Segundo os pesquisadores, nas redes sociais a verdade consome seis vezes mais tempo que uma fake news para atingir 1.500 pessoas. Temos, enfim, a confirmação científica da frase de Mark Twain segundo a qual “uma mentira pode fazer a volta ao mundo no mesmo tempo em que a verdade calça seus sapatos”! (EMPOLI, 2019, p. 46-47)

Acredita-se que essa movimentação social acaba impondo uma “espécie de pacto anticomunitário”, onde uma demanda pela primazia da verdade acaba sendo relativizada, deixando de ser uma “missão prioritária” para determinados grupos sociais, pois estes estão em busca de um local para propagar suas narrativas sem que precise se utilizar de um senso crítico para encontrar a verdade (GIORDAN; MASSI, 2018, p. 283). A conduta é, basicamente, a que Goebbels utilizava na propagação do discurso antissemita, em que a uma mentira repetida mil vezes, em algum momento, acaba se tornando verdade (QUESADA; PISA, 2018, p. 2) e, somado a isso, aproveitam-se do momento crescente de descrença popular com o Estado, as mídias tradicionais, escolas e universidades, para alavancar as “*fake news*, o negacionismo científico, revisionismo histórico e populismo nacionalista” (GIORDAN; MASSI, 2018, p. 284), tendo havido um aumento significativo de *websites* de cunho racista e movimentos nacionalistas brancos, na atualidade, que “chegaram a 14 mil [de *sites* racistas] em 2011 contra (três) em 1995; houve, por sua vez, um aumento de 600% de *sites* nacionalistas brancos”, segundo relatório da ONU (MENDONÇA, 2019, p. 66).

Empoli (2019, p. 49) acredita que a rede social do *Facebook* se tornou uma ferramenta fantástica para multiplicar o discurso de ódio que ultrapassou a barreira virtual e passou a ser parte da vida real. Ele entende que, quando as movimentações surgiram, as demandas eram legítimas como, por exemplo, a indignação sobre aumento das taxas de combustíveis nos Estados Unidos, todavia, o “algoritmo desencadeado pela rede social californiana misturou esses temas com os chamamentos à revolta lançados por extremistas de direita e de esquerda, com *fake news* e com teorias da conspiração de fontes múltiplas”. Sobre os Estados Unidos, ainda, a suposta campanha eleitoral de baixo custo de Donald Trump estava por traz de um grandioso contrato com a empresa de consultoria *Cambridg 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 nação. No caso do Brasil, a campanha de Jair Bolsonaro usou o aplicativo *Whatsapp* para bombardear seus usuários com mensagens carregadas do seu discurso ultranacionalista e notícias falaciosas sobre a oposição. Esse tipo de campanha populista cresceu, demasiadamente, porque “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”. (EMPOLI, 2019, p. 53)

A forma de manipulação através das redes sociais está tão concreta que, segundo os estrategistas da campanha do *Brexit*, no Reino Unido, perceberam que enquanto suas postagens sobre economia rendiam de três ou quatro mil curtidas no máximo, quando era incluído algum “fator emocional” o engajamento aumentava para a casa dos milhares, se dando conta, assim, que “tais sentimentos já estão, de fato, no âmago dessas pessoas”. (EMPOLI, 2019, p. 53) Diante disso, o que nos resta questionar quais seriam as consequências que esses atos refletem na nossa sociedade e, também, para o direito como um todo, visto que, ao que se parece, a utilização das redes sociais na propagação de notícias falsas e do discurso de ódio parecem estar ligadas diretamente aos poderes políticos e econômicos (BRAGA, 2018, p. 215), mas, também, essa capacidade democrática que as redes sociais possuem, é utilizada por agentes não estatais, mais especificadamente, “as comunidades com intenções de prejudicar podem espalhar propaganda e mobilizar seguidores em favor de causas extremistas, como foi visto recentemente com o surgimento do Daesh (Estado Islâmico) e outras organizações terroristas que utilizam as mídias sociais” (SCHWAB, 2016, p. 101).

E como conseguimos enfrentar esses desafios advindos com as novas tecnologias? É primordial que recuperemos valores sociais universais como a dignidade humana, o reconhecimento de bens comuns em detrimento de valores individualistas e a necessidade de mecanismos de gerenciamento do ambiente virtual (SCHWAB, 2019, p. 62), haja vista que a complexidade dos novos fenômenos e seus impactos na sociedade devem ser abordados de forma multilateral e enseja a participação de diversos atores; é necessário a colaboração da área acadêmica, do governo, de líderes empresariais, sociedade civil e organizações internacionais (SCHWAB, 2019. P. 78).

Importante destacar que “a violência verbal jamais irá ser uma via adequada para o enfrentamento de questões polêmicas, de questões necessárias para o desenvolvimento da democracia” ou para encontrar qualquer tipo de solução de situações conflituosas e emblemáticas para nossa sociedade”, pois não é uma forma vantajosa o debate público, possuindo a tendência de prejudicar e romper com sistemas democráticos (AMOSSY, apud, MENDONÇA, 2019, p. 70). Ressalta-se, assim, conforme manifestado pelo documentário *O Dilema das Redes* (2020), que as tecnologias existentes na nossa sociedade pós-moderna em si não são uma ameaça existencial, entretanto, a sua capacidade de trazer à tona o pior que existe nos mais diversos grupos sociais é, de fato, uma ameaça existencial. O que se busca, a partir disso, são soluções para que se possa controlar o que é discutido no âmbito das redes sociais, pois de se um lado temos a

liberdade de expressão assegurada em qualquer regime democrático de direito, por outro lado temos uma enorme quantidade violação de direitos humanos propagados através de *fake news* e do discurso de ódio, e nessa balança de valores e direitos é, extremamente, delicado encontrar um equilíbrio entre essas garantias, trazendo uma enorme responsabilidade, não somente para o legislador, mas também para as empresas privadas em encontrar uma solução rápida e eficiente para esse problema atual.

3. O DIREITO HUMANO À LIBERDADE DE EXPRESÃO E AS INICIATIVAS DE CONTROLE DAS *FAKE NEWS* E DO DISCURSO DE ÓDIO PELAS REDES SOCIAIS

A era da informação, incontestavelmente, revolucionou a sociedade pós-moderna e com ela trouxe diversas inovações e tecnologias, as quais contribuíram para facilitar não somente os afazeres cotidianos, mas também o desenvolvimento tecnológico e científico pelo mundo todo.

No entanto, mesmo que toda essa evolução tenha trazido inúmeros benefícios para a civilização do século XXI, o largo e longo alcance que as redes sociais mostraram serem demasiadamente suscetíveis à disseminação de *fake news* e do discurso de ódio, tendo sido utilizado das mais variadas maneiras para enganar e manipular determinados grupos sociais em prol de uma agenda político-econômica, tendo como ápice as eleições americanas de 2020, quando as redes sociais, como o *Twitter*, começaram a buscar mecanismos para barrar a disseminação desse tipo de informação.

Ocorre que ao se buscar meios para coibir a massificação de *fake news* e o alastramento do discurso de ódio tanto o poder público quanto as plataformas digitais privadas encontram um emblemático problema: como controlar esse tipo de informação nas redes sociais sem que haja violação o direito humano à liberdade de expressão? E para que essa pergunta possa ser respondida, é necessário compreender o que, de fato, é liberdade de expressão e quais são os seus limites dentro do estado democrático de direito e dos direitos humanos, devendo ser visto numa perspectiva mundial.

Primeiramente, muito embora cada país busque conceituar e legislar o direito à liberdade de expressão, não haveria como incluí-los no estudo, tendo em vista a vastidão de legislações relacionadas ao assunto ao redor do mundo. Entretanto, é importante pontuar que no âmbito global podemos nos reger pelo conceito consignado na Declaração Universal dos Direitos Humanos (ORGANIZAÇÃO DAS NAÇÕES UNIDAS, 1948), no qual está disposto em seu artigo XIX que: “toda pessoa tem direito à liberdade de opinião

e expressão; esse direito inclui a liberdade de, sem interferência, ter opiniões e de procurar, receber e transmitir informações e ideias por quaisquer meios e independentemente de fronteiras”.

Ocorre que mesmo definindo o que seria o direito à liberdade de expressão, o DUDH o descreve de forma muito genérica, cabendo inúmeras maneiras de interpretá-lo. Nesse sentido, o Pacto Internacional de Direitos Civis e Políticos (BRASIL, 1992), de 1966, buscou complementar e determinar, em seu art. XIX², de modo mais específico, quais seriam os limites da liberdade de expressão.

O tema é tão relevante e necessário que é preciso que não somente as leis nacionais, mas as de impacto internacionais prescrevam, detalhadamente, quais seriam as formas de liberdade de expressão, bem como seus limites arraigados no princípio basilar da defesa dos direitos humanos, afinal, na era da informação e do imenso desgaste e ofensa aos direitos individuais e coletivos que a massificação das *fake news* e da disseminação do discurso de ódio vem amplificando, cada vez mais com as redes sociais, deverão ser essas as normas que irão nortear a busca por ferramentas que visem coibir os excessos existentes do direito à liberdade de expressão. Ainda, é fundamental destacar, também, que, desde 1969, a Organização dos Estados Americanos (OEA) apresenta na Convenção Americana de Direitos Humanos (CADH) as diretrizes para orientar os limites existentes para a livre expressão individual e coletiva, conforme leciona o art. XIII³.

² Artigo 19. 1. Ninguém poderá ser molestado por suas opiniões. 2. Toda pessoa terá direito à liberdade de expressão; esse direito incluirá a liberdade de procurar, receber e difundir informações e ideias de qualquer natureza, independentemente de considerações de fronteiras, verbalmente ou por escrito, em forma impressa ou artística, ou por qualquer outro meio de sua escolha. 3. O exercício do direito previsto no parágrafo 2 do presente artigo implicará deveres e responsabilidades especiais. Consequentemente, poderá estar sujeito a certas restrições, que devem, entretanto, ser expressamente previstas em lei e que se façam necessárias para: a) assegurar o respeito dos direitos e da reputação das demais pessoas; b) proteger a segurança nacional, a ordem, a saúde ou a moral públicas. (BRASIL, 1992)

³ 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. (ORGANIZAÇÃO DOS ESTADOS AMERICANOS, 1969)

Diante disso, o que se pode observar na legislação internacional é que a liberdade de expressão não se remete apenas a um direito individual, mas também abrange um direito difuso, uma vez que além de conceder o direito de cada cidadão expressar seus sentimentos, ideias, ideais e opiniões de forma livre, permite que esse mesmo indivíduo possa receber qualquer forma de informação, sem que haja qualquer tipo de interferência do poder público ou da iniciativa privada, trazendo a máxima de que “conhecer o pensamento do outro é tão importante quanto exprimir o próprio”. (BENTO, 2016, p. 97)

Não obstante, mesmo a garantia à liberdade de expressão ser um direito fundamental do estado democrático de direito, precisa-se convir que há uma linha tênue entre a liberdade de expressão e a prática do discurso de ódio e a massificação das *fake news*, bem como existe uma tenuidade entre os limites e excessos dessa liberdade às ações ditatoriais. Segundo Bento (2016, p. 98) a proteção à liberdade de expressão “exige um espírito de tolerância e abertura, sem o qual não há uma sociedade pluralista, nem democrática”. Silva (2018, p. 284) acrescenta que ao longo dos anos sempre se pretendeu defender a uma certa tolerância e uma liberdade de expressão mais ampla a partir de grandes tópicos como a religião, imprensa e governo, e, partindo dessa perspectiva, o cidadão estaria demasiadamente prejudicado no que diz respeito a esses temas, caso sua opinião, seja ela qual fosse, tivesse que ser cerceada.

E justo por isso, é primordial reforçar que quando se trata da disseminação do discurso de ódio, seja ele qual for, e do disparo de notícias falsas isso se reflete na “desvalorização do outro como sujeito de direito”, uma vez que há uma vontade de desrespeitar e marginalizar aquilo que lhe é diferente ou a condição para qual contribui para que seja visto como distinto, podendo, inclusive, levar a ofensa ao campo de direitos sociais difusos, os quais pode até ser direcionada a um indivíduo apenas, mas o preconceito e os atributos relacionados ao grupo social ao qual está inserido acaba sendo ferido como um todo. (PAIXÃO; SILVA; CABRAL, 2018, p. 32)

Segundo Paixão, Silva e Cabral (2018, p. 33) para que se possa aferir o nível de ofensa provocado a determinado indivíduo ou a coletividade é primordial que se considere o meio de comunicação empregado, tendo as redes sociais um poder de abrangência e rapidez que facilita o “agrupamento de indivíduos com ideologias semelhantes, potencializa a divulgação de conteúdo, de forma que os ideais viralizam, sendo facilmente aderidos pelos demais, alcançando visibilidade tal que a ideia parece ser

aceita por todos, ou ao menos aceita por todos que realmente importam, revelando-se útil à disseminação de propagandas de ódio”. Pode-se apresentar como exemplo a utilização do ex-presidente Donald Trump que utilizou suas redes sociais para alegar que a eleição americana havia sido fraudada e incitar a população (seus eleitores fiéis) a atacar prédio do Congresso em Washington, durante o qual pessoas foram mortas, tendo o *Facebook*, *Instagram*, *Twitter* e *YouTube* suspendido as contas de Trump devido ao perigo representado por suas mensagens violentas, falsas ou enganosas. O *Twitter* suspendeu contas vinculadas ao grupo *QAnon*, o movimento de extrema direita próximo a Trump. O *Google* e a *Apple* baniram o aplicativo - e a *Amazon.com* parou de hospedar o site - de *Parler*, a plataforma de extrema direita frequentada por seus apoiadores. *Pinterest*, *Reddit*, *Shopify*, *TikTok* e *Twitch* tomaram medidas semelhantes (FLORIDI, 2021).

Diante dessa perspectiva e do uso diário das redes sociais para a disseminação de *fake news* e do discurso de ódio, as empresas das *social medias* passaram a se preocupar com a utilização de ferramentas para essa prática e vêm buscando uma solução para coibir esse tipo de conduta invasiva e prejudicial à sociedade, tendo como recursos inovadores o *The Oversight Board*, do *Facebook*, e a plataforma *Birdwatch*, do *Twitter*.

O projeto *The Oversight Board* foi iniciado em novembro de 2018 e inaugurado em outubro de 2020, tendo como propósito principal auxiliar o *Facebook* a decidir quais conteúdos devem ser removidos da rede, quais serão permitidos e o porquê dessas decisões, partindo da temática da liberdade de expressão. Trata-se de um comitê independente que conta com a participação de, aproximadamente, 40 membros ao redor do mundo e que irão selecionar e julgar conteúdos de postagens já denunciadas no *Facebook* e *Instagram*, decidindo manter ou reverter as decisões das plataformas. (OVERSIGHT BOARD, *Home*)

A ferramenta de julgamento servirá quando o conteúdo do usuário for denunciado e removido pelas redes sociais acima citadas e a plataforma mantiver a decisão após o pedido de análise, baseando-se nas violações dos “Padrões da Comunidade”, podendo, assim, interpor uma apelação ao Comitê de Supervisão, ou *The Oversight Board* (FACEBOOK, Central de ajuda), no prazo de 15 dias, a contar da decisão de remoção de conteúdo. Importante mencionar que o comitê não irá analisar todos os casos recorridos, devendo serem julgados apenas os casos que se qualifiquem como difíceis, significativos, de relevância global e que servirão como base para a criação de políticas futuras. Após a seleção, os componentes do projeto irão deliberar sobre o caso e tomar a decisão baseada em todas as informações fornecidas pelo apelante, pela

rede social e por especialistas, os quais serão chamados para “fornecer contexto”. A decisão será fundamentada e publicada no *site*, bem como terá eficácia vinculante, obrigando a plataforma a implementá-la, exceto se a ação viole a lei (OVERSIGHT BOARD, *Appeals Process*). Desde o início de seus trabalhos, o comitê filtrou milhares de recursos de apelação, tendo julgado e publicado sete casos, os quais já estão disponíveis para leitura e discorrem acerca de discurso religioso, ameaça de violência, desinformação sobre tratamento precoce ineficaz e sem comprovação científica para Covid-19, suposta apologia ao regime nazista, nudez e discurso de ódio (OVERSIGHT BOARD, *Decision*).

Da análise dos julgamentos publicados pelo *The Oversight Board* se pode observar que as decisões do comitê não se restringem apenas às diretrizes e políticas privadas elegidas pelas redes sociais, mas, também, buscam analisar a aplicação da liberdade de expressão a partir de uma perspectiva baseada nos “padrões internacionais de direitos humanos”, tal como descrito na decisão do caso 2020-007-FB-FBR⁴ (OVERSIGHT BOARD, *Decision*).

Ainda assim, o comitê reconhece que é difícil para as plataformas de redes sociais avaliarem a intenção das postagens de cada um de seus usuários e mesmo que a decisão final seja de restaurar o conteúdo removido anteriormente, isso não significa que haja concordância com o conteúdo, mas que, partindo da análise de todo o escopo, não havia sido vislumbrado uma violação direta às leis internacionais que versam sobre os direitos humanos. (OVERSIGHT BOARD, *Decision*. Caso 2020-002-FB-UA)

Em relação ao *Twitter*, a ferramenta criada pela plataforma, para o combate à desinformação pelas *fake news* e a disseminação do discurso de ódio, foi intitulada

⁴ Segundo trecho da decisão do caso 2020-007-FB-FBR, o comitê entendeu que: “de acordo com o Artigo 19º do ICCPR, as pessoas têm o direito de buscar, receber e transmitir ideias e opiniões de todo tipo, inclusive aquelas que possam ser controversas ou profundamente ofensivas (Comentário Geral Nº 34, § 11º). O direito à liberdade de expressão inclui a disseminação de ideias que podem ser consideradas blasfêmias, assim como a oposição a tal discurso. Neste sentido, a liberdade de expressão inclui a liberdade de criticar religiões, doutrinas religiosas e figuras religiosas (Comentário Geral Nº 34, § 48º). A expressão política é particularmente importante e recebe maior proteção sob a lei internacional de direitos humanos (Comentário Geral No. 34, nos §§ 34º e 38º) e inclui apelos a boicotes e críticas a figuras públicas. Ao mesmo tempo, o Comitê reconhece que o direito à liberdade de expressão não é absoluto e pode excepcionalmente estar sujeito a limitações sob as leis internacionais de direitos humanos. Neste caso, após discutir os fatores do Plano de Ação de Rabat, o Conselho não considerou o cargo como sendo de defesa da intolerância religiosa que atinge o limiar do incitamento à discriminação, hostilidade ou violência, que os estados são obrigados a proibir nos termos do artigo 20, parágrafo do § 2º. PIDCP, Artigo 19 (§ 3º exige que as restrições de manifestação sejam facilmente compreendidas e acessíveis (princípio de legalidade), que tenham o propósito de avançar um dos vários objetivos listados (requisito de finalidade legítima), e que sejam necessárias e estreitamente adaptadas ao objetivo específico (princípio de necessidade e proporcionalidade).

Birdwatch e está disponível desde o dia 25 de janeiro de 2021, apenas para os usuários dos Estados Unidos, no qual propaga a intenção de “*Empowering the Twitter community to create a better-informed world*”. Diz, ainda, em um tweet que: “*Want to help build a new community-driven approach to tackling misleading information? Join us — sign up for Birdwatch!*”. (TWITTER, *Birdwatch*).

O objetivo desse recurso, conforme explicado em uma postagem no *blog* do Vice-Presidente de produto do *Twitter*, Keith Coleman (@kcoleman, 2021), é o de expandir além dos rótulos que a empresa já aplica a *tweets* polêmicos ou potencialmente enganosos, que ele sugeriu se limitarem a “*circumstances where something breaks our rules or receives widespread public attention*”. Keith Coleman, ainda, na mesma postagem, relata sobre o piloto nos EUA como uma nova abordagem voltada para a comunidade, isto é, a plataforma contará com a ajudar de seus usuários para lidar com informações enganosas no *Twitter*. A ideia do projeto é a de que as pessoas, ao identificarem informações em *tweets* que acreditam ser enganosas escrevam notas informando. As notas serão, eventualmente, tornadas visíveis o público do *Twitter*, após um consenso amplo e diversificado dos colaboradores. Nesta primeira fase do piloto, as notas serão visíveis apenas em um site *Birdwatch* separado, que poderão os participantes avaliá-las. (COLEMAN, @kcoleman, 2021)

Entretanto, a plataforma do *Twitter* compreende que poderá haver possibilidade de manipulação de tendências, mas que fará todos os esforços possíveis para não permitir a dominação por uma maioria simples, podendo ficar confuso e apresentar as vezes problema, mas que, mesmo assim, investirá neste modelo. Em resumo, a proposta de funcionamento será composta por três elementos: (1) as notas, que são formadas por questões de múltipla escolha e um campo aberto específico, que são públicas – nas quais os participantes adicionam *tweets* que verificarem ser enganosos; (2) as Avaliações, nas quais os participantes avaliam as notas dos outros, na busca de um modelo de reputação futuro de notas mais consistentes; e (3) o site *Birdwatch*, no qual estarão as notas e classificações separados do aplicativo *Twitter*. (COLEMAN, @kcoleman, 2021)

Diante disso, pode-se perceber que as plataformas *Facebook* e *Twitter* estão buscando formas concretas de se combater a desinformação pelas *fake news* e a massificação do discurso de ódio, contudo, as ferramentas de avaliação dos conteúdos postados por seus usuários são, visivelmente, distintas, enquanto o *The Oversight Board* conta com um comitê previamente selecionado e distribuído ao redor do mundo para o julgamento de casos, baseados nas leis internacionais de direitos humanos; o *Birdwatch*

busca trazer essa responsabilidade para a sua comunidade, para que ela, em conjunto, possa criar um modelo ideal de convivência na rede social, criando pressupostos, a partir de questionários, para classificar quais seriam os conteúdos falsos e que violam a dignidade da pessoa humana, para que sejam removidos posteriormente.

CONCLUSÃO

Como se pode observar no decorrer da pesquisa, as redes sociais foram criadas, inicialmente, para que pudesse auxiliar na ampliação do alcance das relações interpessoais nas mais diversas comunidades existentes no mundo, o que não se pode discordar que foi, e ainda é, um grande avanço da sociedade pós-moderna. No entanto, na proporção em que a revolução tecnológica trouxe inúmeros benefícios, ela também vem com um grande potencial de malefícios, diante da rapidez e do alcance e da propagação que a informação possui dentro das redes sociais

No início, partiu-se da hipótese de que a liberdade de expressão embora seja um direito fundamental, amparado não somente pelas legislações existentes nos estados democráticos de direito, mas também nas legislações e diretrizes internacionais, deve haver uma limitação desse direito para que as redes sociais não sejam usadas para a divulgação e propagação irrestrita de *fake news* e do discurso de ódio e as plataformas possam criar ferramentas para coibir e remover esse tipo de conteúdo.

É notório o esforço das mídias sociais de tentar encontrar uma solução para a remoção desses conteúdos nocivos, principalmente depois de terem sido usadas por personalidades de poder para intervir nos mais variados processos democráticos ao redor do mundo, tendo como exemplo as eleições dos Estados Unidos nos anos de 2016 e 2020, bem como nas eleições brasileiras de 2018 e o processo do *Brexit*, no Reino Unido. As inovações vieram com as iniciativas concretas das maiores redes sociais do mundo: o *The Oversight Board*, no *Facebook*; e o *Birdwatch*, no *Twitter*. Enquanto a primeira ferramenta buscou juntar um uma equipe independente ao redor do mundo para que estas pessoas sejam responsáveis para julgar os casos de remoção de postagens da plataforma, baseando-se nas leis internacionais relacionadas ao direito à liberdade de expressão e os limites dispostos nessas diretrizes; a segunda iniciativa apostou no julgamento e boa-fé de sua própria comunidade, com o intuito que todos possam participar na elaboração das medidas atuais e políticas futuras de triagem de conteúdo na rede social. Muito embora o funcionamento das duas iniciativas sejam distintas, o objetivo é o mesmo: a remoção de

conteúdo falso e o discurso de ódio que violem as diretrizes internacionais dos direitos humanos, a fim de garantir o respeito por qualquer pessoas, independentemente de raça, cor, religião ou orientação sexual, bem como a proteção à segurança do estado democrático de direito e tudo que seja inerente a ela.

Todavia, a criação e utilização dessas plataformas de julgamento é muito recente e, apesar de ser uma grande inovação para filtragem de conteúdo postado nas redes sociais, ainda não há como mensurar seu impacto na remoção de postagens falsas e de discurso de ódio nas redes sociais, mas dá para se ter uma ideia da direção que irão caminhar as futuras diretrizes e políticas internas das mídias sociais, uma vez que a cada caso analisado e julgado passarão a moldar e conscientizar a utilização dos aplicativos por seus usuários de uma forma mais cada vez mais consciente e humana.

REFERÊNCIAS

ALVES, Giulia Ferrigno Poli Ide. **Reflexões sobre o fenômeno da desinformação: impactos democráticos e o papel do direito.** Revista dos Estudantes de Direito da Universidade de Brasília – REDUnB. 16º ed. 2019.

BBC NEWS; G1. **Entenda o escândalo de uso político de dados que derrubou valor do Facebook e o colocou na mira das autoridades.** Disponível em: < <https://g1.globo.com/economia/tecnologia/noticia/entenda-o-escandalo-de-uso-politico-de-dados-que-derrubou-valor-do-facebook-e-o-colocou-na-mira-de-autoridades.ghtml> >. Acesso em: 04 abr 2021.

BENTO, Leonardo Valles. **Parâmetros internacionais do direito à liberdade de expressão.** Brasília, 2016.

BRAGA, Renê Moraes da Costa. A indústria das fake news e o discurso de ódio. **Direitos políticos, liberdade de expressão e discurso de ódio: volume I**, 2018.

BRANCO, Sérgio. **Fake news e os caminhos para fora da bolha.** 2017.

BRASIL. Decreto nº 592, de 6 de julho de 1992. **Atos Internacionais. Pacto Internacional sobre Direitos Cíveis e Políticos.** Disponível em: < http://www.planalto.gov.br/ccivil_03/decreto/1990-1994/d0592.htm >. Acesso em: 04 abr 2021.

DEMARTINI, Felipe. **Campanha de Trump usou dados de 50 milhões de usuários do Facebook.** Disponível em: < <https://canaltech.com.br/redes-sociais/campanha-de-trump-usou-dados-de-50-milhoes-de-usuarios-do-facebook-110156/> >. Acesso em: 04 abr 2021.

COLEMAN, Keith. **Introducing Birdwatch, a community-based approach to misinformation.** 25 January 2021. Twitter: @kcoleman. Disponível em: <

https://blog.twitter.com/en_us/topics/product/2021/introducing-birdwatch-a-community-based-approach-to-misinformation.html>. Acesso em: 04 abr 2021.

EMPOLI, Giuliano Da. **Os engenheiros do caos**. São Paulo: Vestígio, 2019.

FACEBOOK, **Central de ajuda**. Disponível em: <<https://www.facebook.com/help/346366453115924>>. Acesso em: 04 abr 2021.

FLORIDI, Luciano. **Las redes sociales no se pueden regular solas**. EL PAIS, 2021. Disponível em: <<https://elpais.com/opinion/2021-01-25/las-redes-sociales-no-se-pueden-regular-solas.html>>. Acesso em: 04 abr 2021.

FORNASIER, Mateus de Oliveira; BECK, Cesar. (2020). **Cambridge analytica: escândalo, legado e possíveis futuros para a democracia**. *Revista Direito Em Debate*, 29(53), 182-195. Disponível em: <<https://doi.org/10.21527/2176-6622.2020.53.182-195>>. Acesso em: 04 abr 2021.

GIORDAN, Marcelo; MASSI, Luciana. **Formação e atuação do tutor como orientador de pesquisa na educação on-line**. ETD - Educação Temática Digital, Campinas, SP, v. 20, n. 2, p. 495–517, 2018. DOI: 10.20396/etd.v20i1.8646237. Disponível em: <<https://periodicos.sbu.unicamp.br/ojs/index.php/etd/article/view/8646237>>. Acesso em: 04 abr 2021.

JUNIOR, Demercino. **A Propaganda Nazista**. Disponível em: <<https://brasilecola.uol.com.br/historiag/propagandanazista.htm>>. Acesso em: 04 abr 2021.

MCINTYRE, Lee. **Pós-verdade**. Imprensa, 2018, p. 1.

MENDONÇA, Analméria da Silva Cabral de. **Liberdade de expressão nas mídias sociais: discursos de ódio e notícias falsas como meio de violação dos Direitos Humanos nas interlocuções virtuais**. Recife, 2019.

O DILEMA DAS REDES. Direção de Jeff Orlowski. Estados Unidos: Netflix, 2020.

OLIVEIRA, André Soares; GOMES, Patrícia Oliveira. Os limites da liberdade de expressão: fake news como ameaça à democracia. **Revista de Direitos e Garantias Fundamentais**, v. 20, n. 2, 2019.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS. **Declaração Universal dos Direitos Humanos**, 1948. Disponível em: <<https://www.unicef.org/brazil/declaracao-universal-dos-direitos-humanos>>. Acesso em: 04 abr 2021.

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: 04 abr. 2021.

OVERSIGHT BOARD. **Appeals Process**. Disponível em: <<https://oversightboard.com/appeals-process/>>. Acesso em: 04 abr 2021.

OVERSIGHT BOARD. **Case Decision 2020-002-FB-UA**. Disponível em: < <https://oversightboard.com/decision/FB-I2T6526K/>> Acesso em: 04 abr 2021.

OVERSIGHT BOARD. **Case Decision 2020-007-FB-FBR**. Disponível em: < <https://oversightboard.com/decision/FB-R9K87402/>> Acesso em: 04 abr 2021.

OVERSIGHT BOARD. **Decision**. Disponível em: < <https://oversightboard.com/decision/>>. Acesso em: 04 abr 2021.

OVERSIGHT BOARD. **Home**. Disponível em: < <https://oversightboard.com/> >. Acesso em: 04 abr 2021.

PAIXÃO, Alessandro Gonçalves da; SILVA, Débora Pereira; CABRAL, Nuria Micheline Meneses **Liberdade de expressão e hate speech no estado democrático de direito**. Revista de Direito, [S. l.], v. 10, n. 01, p. 23, 2018. DOI: 10.32361/201810011478. Disponível em: < <https://periodicos.ufv.br/revistadir/article/view/1478> >. Acesso em: 5 abr. 2021.

QUESSADA, Miguel; PISA, Licia Frezza. **Fake News Versus MIL: a difícil tarefa de desmentir Goebbels**. Belo Horizonte, 2018.

SANTOS, Arantxa Carla da Silva; PAES, Renata da Cruz; PONTES, Altem Nascimento. **Mídia pós-massiva: um levantamento de podcast especializado em meio ambiente como instrumento de conscientização ambiental**. Texto Livre: Linguagem e Tecnologia, Belo Horizonte-MG, v. 12, n. 1, p. 153–168, 2019. DOI: 10.17851/1983-3652.12.1.153-168. Disponível em: < <https://periodicos.ufmg.br/index.php/textolivre/article/view/16838> >. Acesso em: 04 abr 2021.

SCHWAB, Klaus. **A quarta revolução industrial**; tradução Daniel Moreira Miranda. São Paulo: Edipro, 2016.

SCHWAB, Klaus; DAVIS, Nicholas. **Aplicando a quarta revolução industrial**. Edipro, 2019

SILVA, Peterson Roberto da. **O conceito de “Liberdade de expressão”**, Florianópolis, 2018.

TWITTER. **Birdwatch**. Disponível em: < <https://twitter.com/birdwatch> >. Acesso em: 04 abr 2021.

WILKE, Valéria Cristina Lopes. **Pós-verdade, fake News e outras drogas, vivendo em tempos de informação tóxica**. Logeion: Filosofia da Informação, v. 7, p. 8-27, 2020.