

**MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DO PIAUÍ - UFPI
PRÓ-REITORIA DE ENSINO DE PÓS-GRADUAÇÃO
CENTRO DE CIÊNCIAS HUMANAS E LETRAS - CCHL
PROGRAMA DE PÓS-GRADUAÇÃO EM DIREITO - PPGD**

REBEKA COELHO DE OLIVEIRA CARVALHO

**MILÍCIAS DIGITAIS E FILTRAGEM POR INTELIGÊNCIA ARTIFICIAL: uma
análise constitucional das exclusões de fake news disseminadas por milícias digitais pelas
empresas gestoras dos fluxos comunicacionais online**

TERESINA - PI
2023

REBEKA COELHO DE OLIVEIRA CARVALHO

**MILÍCIAS DIGITAIS E FILTRAGEM POR INTELIGÊNCIA ARTIFICIAL: uma
análise constitucional das exclusões de fake news disseminadas por milícias digitais pelas
empresas gestoras dos fluxos comunicacionais online**

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Direito, da Universidade Federal do Piauí, como parte dos requisitos para a obtenção do título de Mestra em Direito.

Área de concentração: Direito, Democracia e Mudanças Institucionais

Linha de Pesquisa: Mudanças Institucionais, Constituição, Democracia e Justiça

ORIENTADOR: Prof. Dr. Sebastião Patrício Mendes Costa

FICHA CATALOGRÁFICA

Universidade Federal do Piauí
Biblioteca Comunitária Jornalista Carlos Castello Branco Divisão de Representação da
Informação

C331m Carvalho, Rebeka Coelho de Oliveira.

Milícias digitais e filtragem por inteligência artificial : uma análise constitucional das exclusões de fake news disseminadas por milícias digitais pelas empresas gestoras dos fluxos comunicacionais online / Rebeka Coelho de Oliveira Carvalho. -- 2023.

194 f.

Dissertação (Mestrado) – Universidade Federal do Piauí, Centro de Ciências Humanas e Letras, Programa de Pós-Graduação em Direito, Teresina, 2023.

“Orientador: Prof. Dr. Sebastião Patrício Mendes Costa”.

1. Liberdade de expressão. 2. Fake news. 3. Milícias digitais.
4. Filtragem algorítmica. I. Costa, Sebastião Patrício Mendes.
II. Título.

CDD 341.2732

Bibliotecária: Francisca das Chagas Dias Leite – CRB3/1004

REBEKA COELHO DE OLIVEIRA CARVALHO

MILÍCIAS DIGITAIS E FILTRAGEM POR INTELIGÊNCIA ARTIFICIAL: uma análise constitucional das exclusões de fake news disseminadas por milícias digitais pelas empresas gestoras dos fluxos comunicacionais online

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Direito, da Universidade Federal do Piauí, como parte dos requisitos para a obtenção do título de Mestra em Direito.

Área de concentração: Direito, Democracia e Mudanças Institucionais

Linha de Pesquisa: Mudanças Institucionais, Constituição, Democracia e Justiça

ORIENTADOR: Prof. Dr. Sebastião Patrício Mendes Costa

APROVADA EM: 27 de fevereiro de 2023.

BANCA EXAMINADORA

Prof. Dr. Sebastião Patrício Mendes Costa (UFPI)
Orientador

Prof^a. Dr^a. Olívia Brandão Melo Campelo (UFPI)
Examinadora Interno

Prof^a. Dr^a. Gabrielle Bezerra Sales Sarlet (PUC-RS)
Examinadora Externa

Teresina - PI
2023

AGRADECIMENTOS

A Vinicius Carvalho meu amor, marido, parceiro e incentivador! Muito obrigada por dividir esse momento comigo e tornar possível a conclusão do mestrado. Você é muito importante em minha vida! Você é incrível! Amo você!

Aos meus filhos, Valentina e Vinicius, obrigada pela motivação! Apesar de ainda serem crianças, espero que tenham compreendido minhas ausências e meus pensamentos distantes, processando as leituras. Como sou feliz por ser a mãe de vocês, quanto orgulho tenho dos dois! E nesses dois anos de mestrado, ambos me disseram expressamente: “mamãe tenho orgulho de você!” E isso ficará para sempre comigo! Cada dia procuro ser uma pessoa melhor, para ser uma mãe melhor! Amo infinitamente vocês, Valentina e Vinicius Cosme!

Aos meus pais que estão imensamente felizes com minha realização pessoal!

A minha família, irmãos, cunhados, sobrinhos e sogra que me ajudaram e me apoiaram muito nessa jornada.

Ao meu orientador, Professor Doutor Sebastião Costa, que foi paciente, perseverante e procurou sempre passar tranquilidade diante da minha intensidade. Muito obrigada!

A professora Doutora Olívia Brandão Melo Campelo que foi uma grande inspiração!

RESUMO

A comunicação social online impacta fortemente a democracia. A estruturação e organização das informações online por empresas privadas que fazem a gestão dos fluxos comunicacionais perseguindo lucros tem causado efeitos na interação entre os cidadãos, na formação da opinião pública e influenciado escolhas político-eleitorais. A internet se tornou um modelo de negócios, por isso, conteúdos que prolongam a navegação online e causam maior agitação são priorizados e assim, as fake news adquiriram maior destaque que informações verificadas e balanceadas. Como a internet assumiu uma onipresença na vida das pessoas, passou também a ser a principal fonte de informações e do mesmo modo que viabilizou a pluralidade fontes de informação e possibilitou uma democratização de acesso às mídias também abriu espaço para disseminação de fake news e para atividades fraudulentas e, nesse aspecto, as milícias digitais desequilibram as regras do jogo democrático. Diante da incapacidade técnica do Estado para fazer o controle de conteúdo na internet, a infraestrutura privada gestora dos fluxos comunicacionais online tem sido pressionada para desempenhar essa função, no entanto, em razão do volume de dados que circulam online o uso de tecnologia se tornou indispensável para monitorar e controlar o trânsito de informações. O estudo analisa a constitucionalidade das filtragens algorítmicas de fake news disseminadas por milícias digitais com o uso de inteligência artificial operadas pelas empresas gestoras dos fluxos comunicacionais. Para tanto, traça um panorama da liberdade de expressão na internet e como o modelo de negócios online maneja esse direito segundo finalidades lucrativas. Assim, apesar de se influenciarem mutuamente os meios digital e físico não se equivalem, razão pela qual foram diferenciados esfera pública e ambiente de discussão online, participação online de expressão cívica numa democracia direta, além de apontado o contraste da internet que, apesar de ser um ambiente que contém informações balanceadas, verificadas e de fontes confiáveis, não as privilegia. Foram caracterizados os efeitos das análises algorítmicas como normas de comando geral, assim como demonstrado que não há uniformidade entre os critérios de moderação de conteúdo nas redes. Também se investiga como a liberdade de expressão e as informações têm sido manejadas segundo um projeto de instrumentalização da humanidade além de diferenciada a governança da internet e a democracia. O terceiro capítulo trata da desinformação online, caracteriza as fake news, analisa o controle de conteúdo caso a caso, descreve a atividade miliciana e o potencial ainda mais lesivo do comportamento articulado em rede porque, mais do que a mera divulgação de fake news, as milícias, além de destruírem a realidade compartilhada, muitas vezes têm o objetivo específico de criar artificialmente falsos cenários e, por último, explica como a filtragem algorítmica pode ser um mecanismo eficiente para contingenciamento célere de fake news. O quarto capítulo contextualiza o enfrentamento das fake news no Brasil, especialmente, nos processos eleitorais e a tendência do controle sistêmico. Ao fim, analisa a constitucionalidade da exclusão algorítmica de fake news disseminadas por milícias digitais operadas pelas empresas gestoras dos fluxos comunicacionais com o uso de inteligência artificial, por qualificar as fake news disseminadas por milícias digitais como elemento subversivo às regras do jogo democrático.

Palavras-chaves: Fake news; liberdade de expressão; milícias digitais; filtragem algorítmica.

ABSTRACT

The online media strongly impacts democracy. The structuring and organization of online information by private companies that manage communication flows in pursuit of profits has had effects on the interaction between citizens, on the formation of public opinion and has influenced political-electoral choices. The internet has become a business model, so content that prolongs online browsing and causes more excitement is prioritized, and thus, fake news has gained greater prominence than verified and balanced information. As the internet assumed omnipresence in people's lives, it also became the main source of information and, in the same way that it made possible the plurality of information sources and enabled a democratization of access to the media, it also opened space for the dissemination of fake news and for activities fraudulent and, in this respect, the digital militias unbalance the rules of the democratic game. Faced with the technical inability of the State to control content on the internet, the private infrastructure that manages online communication flows has been pressured to perform this function, however, due to the volume of data circulating online, the use of technology has become indispensable to monitor and control the flow of information. The study analyzes the constitutionality of algorithmic filtering of fake news disseminated by digital militias with the use of artificial intelligence operated by companies that manage communication flows. To do so, it outlines an overview of freedom of expression on the internet and how the online business model manages this right according to profitable purposes. Thus, despite mutually influencing the digital and physical media, they are not equivalent, which is why the public sphere and online discussion environment were differentiated, online participation of civic expression in a direct democracy, in addition to pointing out the contrast of the internet which, despite being an environment that contains balanced, verified information from reliable sources does not privilege it. The effects of algorithmic analyzes were characterized as general command norms, as well demonstrated that there is no uniformity between the content moderation criteria in the networks. It was also investigated how freedom of expression and information have been managed according to a project of instrumentalization of humanity in addition to differentiated internet governance and democracy. The third chapter dealt with online misinformation, characterized fake news, analyzed content control on a case-by-case basis, described militia activity and the even more harmful potential of networked behavior because, more than the mere dissemination of fake news, the militias, in addition to destroying shared reality, often have the specific objective of artificially creating false scenarios and, finally, explained how algorithmic filtering can be an efficient mechanism for the rapid contingency of fake news. The fourth chapter contextualizes the confrontation of fake news in Brazil, especially in electoral processes and the tendency of systemic control. Finally, it analyzes the constitutionality of the algorithmic exclusion of fake news disseminated by digital militias operated by companies that manage communication flows with the use of artificial intelligence, by qualifying fake news disseminated by digital militias as a subversive element to the rules of the democratic game.

Keywords: fake news; freedom of expression; digital militias; algorithmic filtering

SUMÁRIO

1 INTRODUÇÃO	8
2 LIBERDADE DE EXPRESSÃO E INTERNET	16
2.1 Ambiente de discussão online	20
2.1.1 Modelo de negócios.....	20
2.1.2 Ciberativismo	27
2.2 Comunicação social eletrônica (Art. 222 § 3º da CF/88)	31
2.2.1 Informação em larga escala.....	35
2.2.2 Algoritmos e grande volume de dados	40
2.3 Moderação de conteúdo, <i>lato sensu</i>, no ciberespaço e o desafio de fazer respeitar direitos fundamentais no ambiente virtual	47
2.3.1 Expansão dos critérios de modo discricionário sem uma diretiva regulatória específica.....	48
2.3.2 Liberdade de expressão no ambiente online e instrumentalização	53
2.4 Democracia e governança das redes	59
2.4.1 Governança das redes, liberdade de expressão e regulação	59
2.4.2 Diferença entre a governança da internet e a democracia considerando as características procedimentais de Bobbio	63
3 DESINFORMAÇÃO ONLINE	69
3.1 Fake news	70
3.2 Fake news e a insuficiência / problemática do modelo de Moderação Online Judicial	81
3.3 Milícias digitais	93
4 INFLUÊNCIAS INTERNACIONAIS JURISPRUDENCIAIS E LEGISLATIVAS DA DINÂMICA VIRTUAL NO BRASIL E COMO AS FORMAS DE COMBATE ÀS FAKE NEWS TÊM EVOLUIDO NO AMBITO DO PODER JUDICIÁRIO, ESPECIALMENTE NO TSE	108
5 ANÁLISE CONSTITUCIONAL DA FILTRAGEM ALGORÍTMICA DE FAKE NEWS DISSEMINADAS POR MILÍCIAS DIGITAIS	136
6 CONCLUSÃO	171
REFERÊNCIAS	182

1 INTRODUÇÃO

A internet trouxe novos horizontes para a liberdade de expressão e desencadeou consequências até então inéditas em razão das novas dimensões virtuais que alteraram a percepção de espaço, velocidade e tempo, aumentando a amplitude de disseminação e a velocidade de propagação, ultrapassando fronteiras em instantes e inaugurando uma nova forma de comunicação social, com interações recíprocas e síncronas.

Apesar da internet ter possibilitado uma variedade e mesmo multiplicidade de fontes de informação, essenciais ao regime democrático, a democratização de acesso ao público trouxe implicações graves à modernidade dentre quais: as fake news. Contornar esse problema apresenta um desafio, tendo em vista que democracia pressupõe liberdade de expressão com debate de ideais para escolhas públicas.

Diante disso, percebe-se que a revolução tecnológica provocou alterações nas relações sociais ao democratizar o acesso às mídias, sendo os usuários os próprios produtores e consumidores de informações, mas também criou riscos para a democracia e novos espaços para atividades fraudulentas. Como em regra não há checagem prévia de conteúdo antes da publicação online, as fake news se multiplicaram no ambiente virtual inclusive com disparo maciço por milícias digitais¹.

As milícias digitais são comportamentos articulados em rede que usam a lógica algorítmica e podem ter uma intenção determinada, sendo atividade não apenas para tumultuar, gerar lucro ou prender a atenção da audiência. A ação milicianiana coordenada muitas vezes tem um objetivo definido, de influenciar a realidade compartilhada segundo uma direção específica.

As fake news têm como característica provocar prejuízos ou aferir ganhos e as milícias digitais potencializam esses efeitos ao promover um ataque maciço e articulado, acionando os mecanismos de impulsionamento algorítmicos em rede e expandindo o espectro de usuários atingidos, de modo a deformar a avaliação da realidade compartilhada, comprometer a formação da opinião pública e, por consequência, influenciar as escolhas democráticas.

A interconectividade entre os usuários², através das trocas de mensagens, produz uma quantidade de dados acima da cognição humana de processamento razão pela qual se tornou

¹ Entende-se aqui por milícias digitais as organizações virtuais que geram falsas informações a fazem circular na internet, em redes sociais e demais aplicativos, gerando impactos sociais e políticos significativos.

² Neste trabalho, tem-se por usuário a pessoa que utiliza o sistema inteligente.

necessário organizar e estruturar as informações online, estabelecendo prioridades dos fluxos mediante análises algorítmicas. A comunicação social online tornou imprescindível a intermediação das informações por grandes empresas gestoras dos fluxos comunicacionais de modo a facilitar as pesquisas e os encontros virtuais de contatos próximos ou com preferência comuns.

Os contratos e termos de serviços firmados por essas empresas com milhões de usuários individuais contêm cláusulas que lhes favorecem, possuindo muitos direitos e poucas responsabilidades. Desse modo, permite-se a expansão do poder continuamente, seja ao se apoderarem de dados pessoais, ao controlarem o acesso às suas comunidades e à comunicação online, quanto pela revisão contínua de suas cláusulas, criando critérios de moderação de conteúdo que repercutem em restrições cada vez maiores sobre a liberdade de expressão dentro de seus espaços, mercantilizando esse direito. Nesta feita, todo o fluxo online passa a ser gerido para manter as pessoas em moldes privados, instrumentalizando suas vontades, restringindo as oportunidades de fala aos interesses lucrativos e não mais no interesse da comunicação social.

Como os fluxos comunicacionais online são geridos por empresas que se guiam por finalidades lucrativas, a governança digital privilegia conteúdo que promove a agitação online e engajamento, como guerras de narrativas com compartilhamentos, comentários e curtidas e por essa razão, privilegia as fake news. Além disso, as notícias e informações online, mesmo as verificadas, são moduladas, ou seja, filtradas e direcionadas, segundo as preferências do perfil do usuário, reforçando convicções, radicalizando opiniões e dificultando o diálogo colaborativo, próprio da democracia. Assim, a finalidade lucrativa do ambiente online, ao moderar e modular conteúdo acabou por dividir os cidadãos e distanciar os opostos, dificultando o entendimento entre pessoas com posições diferentes.

Nesse contexto, a internet não se guia por interesses públicos nem representa fielmente a realidade física e fática e, apesar de muitas atividades cotidianas terem uma versão correspondente no espaço virtual, estas não se equivalem. Assim, o ambiente de discussão online não corresponde à esfera pública, porque é um modelo de negócios projetado artificialmente para atrair atenção do público e direcionar conteúdo segundo finalidades lucrativas cujo propósito não é o entendimento, e sim a polarização e a conflituosidade.

Da mesma forma, ciberativismo não é o mesmo que cidadania, debate democrático não é um processo que possa ser internalizado mediante opções de enquete, como cliques de sim ou não, tendo em vista que as manifestações online são aleatórias e não há procedimentos prévios definidos com trâmites e prazos. Nesse sentido, participação interativa entre os usuários, com

respostas binárias, emojis, argumentos baseados em emoções não são expressão de democracia direta.

Ocorre que a maior parte dos cidadãos se informa através da internet, congestionada por fake news, e apesar do ambiente de discussão online ter potencial para ser uma fonte confiável e plural, os fluxos online não priorizam informações balanceadas ou verificadas, sendo essa a razão pela qual os fluxos comunicacionais online influenciam tanto a democracia e as escolhas públicas.

Os constantes avanços tecnológicos associados às dinâmicas online que remuneram os produtores de conteúdo de acordo com o engajamento que provocam online tem incrementado a produção e distribuição de fake news, inclusive mediante ação articulada em rede como é o caso das milícias digitais. Se a produção e distribuição das fake news se aperfeiçoaram pelo desenvolvimento tecnológico, o controle caso a caso tem se mostrado cada vez mais problemático e a situação é ainda mais complexa quando a disseminação envolve atividade de milícias digitais que extrapolam as jurisdições nacionais.

O processo de digitalização da atividade humana está cada vez mais acelerado e isso repercute em muitos dados online. O trânsito desses dados tornou a gestão tecnológica dos fluxos uma medida essencial. Assim, análises algorítmicas com o uso de inteligência artificial³ capazes de processar grande quantidade de dados, de maneira rápida e dinâmica, têm cada vez mais relevância numa sociedade de consumo online em massa.

Os critérios sobre as exclusões de conteúdo desinformativo online, operados de maneira automatizada, passaram a ser uma questão central para a democracia tendo em vista que as restrições arbitrárias à liberdade de expressão online mercantilizam esse direito e instrumentalizam a vontade dos usuários. De igual modo, a livre disseminação de fake news, sem contingenciamento, acaba por afetar direitos da personalidade e bens coletivos juridicamente relevantes.

A filtragem algorítmica de fake news com o uso de inteligência artificial pode interromper a distribuição e a consecução dos fins fraudulentos das milícias digitais e se apresenta como uma ferramenta tecnológica rápida, apesar de não ser tão precisa, para contingenciar, impedir ou minimizar os efeitos das fake news online disseminadas por milícias

³ De acordo com os autores Fabro Steibel, Victor Freitas Vicente e Diego Santos Vieira de Jesus (2019) inteligência artificial “refere-se à habilidade de um sistema interpretar corretamente dados externos, aprender a partir desses dados e usar o aprendizado para traçar objetivos e tarefas específicos por meio de adaptação flexível.”

digitais. Assim como frear a remuneração pela circulação das fake news e trazer a tona informações e notícias verificadas e balanceadas, produzidas segundo padrões éticos e profissionais.

E nesse aspecto, a moderação de conteúdo pelas empresas gestoras dos fluxos comunicacionais se mostra fundamental porque o Estado, em virtude da incapacidade técnica, teve o papel diminuído no controle de conteúdo online e passou a pressionar a infraestrutura privada da internet, que faz a gestão dos fluxos comunicacionais online globais, a exercerem essa função.

A comunicação social eletrônica online em sociedades de massa conectada e de intensa participação nas redes, necessita de um modelo de controle genérico que não dependa somente da análise circunstanciada caso a caso, com mecanismos de contingenciamento do fluxo comunicacional online mais compatíveis com a velocidade de propagação do meio virtual. No Brasil, para os usuários, o modelo regulatório prioriza o controle jurisdicional, no entanto, os trâmites e formalidades do Poder Judiciário se mostram em descompasso com o dinamismo online tendo em vista a quantidade de manifestações online diárias.

A moderação de conteúdo no Brasil tem como padrão a análise caso a caso, através da via judicial, mas o controle de fake news online não tem correspondido às necessidades de contingenciamento célere tendo em vista que os julgamentos individualizados, sejam judiciais ou internos, demandam tempo e as fake news se espalham velozmente. Além disso, enquanto não excluídas, continuam a projetar efeitos que são difíceis de mensurar porque o caminho percorrido em rede é imprevisível e desconhecido, prejudicando até mesmo o direito de resposta e a aferição dos danos.

E mais, os meios constitucionais de reparação pelo exercício abusivo da liberdade de expressão, como direito de resposta e indenização pecuniária, são insuficientes porque os perpetradores de fake news muitas vezes são anônimos ou difíceis de serem identificados, podendo ser disseminadores artificiais ou contas inautênticas. Mesmo na circunstância de serem localizados, muitas vezes não têm condições de arcar financeiramente com o valor das indenizações. Some-se ainda que, quanto mais se tenta revelar a verdade, no caso do direito de resposta, mais camadas de fake news são lançadas, tendo em vista que o retorno do assunto só reacende as discussões. No âmbito online, a confiabilidade e a verificação dos fatos têm sido relegadas em detrimento de convicções pessoais e opiniões que mais se adaptam ao interesse do usuário, ainda que não tenham correspondência com a realidade fática.

Como não há lei que determine o que pode ou não ser dito, vigorando o princípio da legalidade do artigo 5º, II, da CF/88, o Supremo Tribunal Federal (STF) é que tem construído os limites da liberdade de expressão. O inquérito das fake news, de 2020, que tramita em segredo de justiça, pode trazer novos elementos que caracterizam o exercício arbitrário da liberdade de expressão, assim como o julgamento de ações pendentes em controle concentrado de constitucionalidade podem trazer luzes sobre qual o entendimento da corte diante das dinâmicas virtuais e novas formas de contingenciamento.

Nesse sentido, a liberdade de expressão tem sido ressignificada com seus contornos, moldados segundo as novas dimensões virtuais, do mesmo modo as formas de controle tradicionais, mediadas pelo Poder Judiciário, começam a atuar em parcerias com as empresas gestoras dos fluxos comunicacionais. É um caminho de aprendizagem, a ser calibrado de acordo com as experiências, mas inevitável diante do cenário mundial conectado em que a liberdade de expressão migra do espaço público para o ambiente privado e do controle estatal para o global.

No Brasil, o combate a atividade miliciana com disparo de fake news tem sido mais intenso no período eleitoral e para dar maior celeridade ao enfrentamento desse comportamento articulado foi editada a Resolução 23.714 em outubro 2022, fixando multas vultosas em caso de descumprimento de determinação judicial com prazo contado de horas, dando poderes ao Presidente do Tribunal Superior Eleitoral (TSE) para estender os efeitos de decisões do Tribunal sobre exclusões de conteúdo desinformativo relacionadas ao processo eleitoral brasileiro, com identidade de conteúdo, para outros endereços eletrônicos e para determinar a suspensão de serviços das plataformas.

A questão é encontrar o ponto de equilíbrio entre a avaliação de perdas e ganhos feitas pelas empresas e a prevalência do interesse público perante o Estado. Na relação liberdade de expressão online e democracia, a finalidade lucrativa da estruturação e organização dos fluxos comunicacionais concorre com os direitos à informação, ao pluralismo político e à cidadania. No tocante ao contingenciamento da liberdade de expressão, a avaliação segundo custos e benefícios quanto às exclusões de fake news feita pelas próprias empresas em seus domínios compete com a apreciação de perigo real e iminente no âmbito judicial segundo o modelo padrão de moderação de conteúdo online.

Assim, a influência da doutrina e jurisprudência americanas sobre a liberdade de expressão e as construções europeias sobre modelos de controle das práticas online mediante a autorregulação regulada e o princípio da autodeterminação informativa no intuito de preservar

direitos de terceiros nas redes tem contribuído para um aprofundamento do combate às fake news e o aperfeiçoamento dos mecanismos de controle.

A ideia central da pesquisa é que a mesma tecnologia de inteligência artificial⁴ que estrutura a comunicação, pode ser usada para realizar controle de conteúdo desinformativo, procedendo varreduras sistêmicas nos domínios virtuais, detectando e excluindo fake news disseminadas em larga escala, no intuito de viabilizar um ambiente de discussão menos artificial e mais próximo da realidade fática. Para tanto, essas exclusões têm de ser equilibradas, sem favorecimento a interesses ideológicos específicos. A motivação para a pesquisa foi encontrar um critério para exclusão de fake news de modo sistêmico sem depender de denúncias prévias ou ordem judicial anterior, promovendo desta forma uma proteção mais efetiva a democracia e seus institutos e direitos constitucionalmente tutelados.

Tornar o ambiente de discussão online menos poluído por fake news, sem comprometer a liberdade de expressão, pode contribuir para melhorar a qualidade da informação, a interação entre usuários com opiniões divergentes além de facilitar a construção de posições conciliadas numa sociedade polarizada. Assim, o estudo tem como finalidade responder a seguinte indagação: a exclusão algorítmica⁵ de fake news disseminadas por milícias digitais operada pelas empresas gestoras dos fluxos comunicacionais configura restrição indevida à liberdade de expressão?

Desse modo, a pesquisa tem como objetivo analisar se a exclusão algorítmica de fake news disparadas por milícias digitais é constitucional ou não. Nesses termos, visa legitimar um critério para moderação online em que o controle de conteúdo e de padrão de comportamento online se somam para identificar e remover fake news disseminadas por milícias digitais, mediante filtragens algorítmicas com o uso de inteligência artificial, podendo esse contingenciamento ser operado de modo proativo pelas empresas gestoras dos fluxos da comunicacionais online, dentro de seus domínios, no intuito de diminuir a circulação de conteúdo desinformativo.

⁴ Aqui tido por “Modelo de Inteligência Artificial: conjunto de dados e algoritmos computacionais, concebidos a partir de modelos matemáticos, cujo objetivo é oferecer resultados inteligentes, associados ou comparáveis a determinados aspectos do pensamento, do saber ou da atividade humana.” (Art. 3º, II, da Resolução 332 do CNJ de 2020)

⁵ Para melhor compreensão tem-se algoritmo como expressão matemática utilizada como instrução para a resolução de um problema. Para Pitruzzella (2018) algoritmo é um procedimento de cálculo matemático, que pode ser descrito com um número finito de regras, ou seja, uma instrução processualizada para a execução de uma operação mais ou menos complexa ou a resolução de um problema.

Como objetivos específicos visa diferenciar o ambiente de discussão online da esfera pública segundo o conceito teórico de Habermas, distinguir participação online de expressão cívica numa democracia direta, apontar o contraste da internet que, apesar de ser um ambiente que contém informações balanceadas, verificadas e de fontes confiáveis, não as prioriza. Também objetiva caracterizar os efeitos das análises algorítmicas como normas de comando geral, demonstrar que não há uniformidade nos critérios de moderação de conteúdo entre as redes, investigar como a liberdade de expressão e as informações têm sido manejadas segundo um projeto de instrumentalização da humanidade e apontar a necessidade de proteção sistêmica contra as fake news.

Além disso, diferencia a governança da internet da democracia, segundo as características procedimentais da teoria de Bobbio, caracteriza as fake news abordando contexto e conteúdo e aspectos como finalidades, motivos pelos quais prosperam e efeitos, bem como analisa como o controle de conteúdo caso a caso é ineficiente diante do volume de dados online e da necessidade de medidas céleres de contingenciamento. De igual modo, descreve a atividade miliciana e o potencial ainda mais lesivo do comportamento articulado em rede⁶ porque mais do que a vantagem ou o prejuízo a terceiro, as milícias, além de destruírem a realidade compartilhada, muitas vezes têm o objetivo específico de criar artificialmente falsos cenários.

Explica-se ainda como a filtragem algorítmica apesar de apresentar falhas, pode ser uma ferramenta célere de contingenciamento de fake news para impedir ou minimizar os efeitos danosos a direitos da personalidade e bens coletivos juridicamente relevantes como a democracia. Por fim, correlaciona as influências doutrinárias e jurisprudenciais internacionais que influenciaram a compreensão da liberdade de expressão no país, bem como demonstrou a mudança da atuação estatal que passou de uma postura de não intervenção para formas de controle com parcerias público-privadas como na atuação do TSE no combate as fake news.

Este estudo decorre da necessidade social de trazer maior segurança e confiabilidade às informações que são veiculadas no ambiente de discussão online, diminuindo a propagação de fake news. Em razão da intensidade das interações no meio virtual, do grande volume de dados que circula nas redes e dos dilemas que envolvem controle da liberdade de expressão no contexto democrático, a importância científica dá-se no sentido de aprofundar a compreensão

⁶ Utiliza-se a seguinte compreensão: “rede é um conjunto de nós interconectados. Nó é o ponto no qual uma curva se entrecorta”. (CASTELLS, 2020, p. 553/554)

da liberdade de expressão no ambiente online estudando a possibilidade de contingenciamento sistêmico das fake news disseminadas por milícias digitais.

Assim, a pesquisa se guiou pelo método dedutivo, com uma natureza qualitativa com objetivo exploratório e utilizou como fonte material bibliográfico como publicações, dados coletados que se encontram organizados em arquivos, banco de dados, anuários, estatísticas, relatórios, doutrina, artigos científicos, legislação nacional, estrangeira, reportagens jornalísticas, decisões jurisprudenciais e projetos de lei.

A estrutura do trabalho se estabelece da forma a seguir descrita. Em um primeiro momento, descreve-se o universo a ser pesquisado, que é o ambiente online em que as fake news disseminadas por milícias digitais online se constituem forma de desinformação e se revelaram como elemento subversivo às regras do jogo democrático. Os modelos de controle fora do âmbito jurisdicional enfrentam críticas quanto ao cerceamento da liberdade expressão e o comprometimento da pluralidade democrática tendo em vista o risco de silenciamento de vozes divergentes ou de minorias tão essenciais à democracia, por isso os valores da liberdade de expressão de Jack Balkin (2020) nortearam a compreensão desse direito e deram suporte as conclusões.

O segundo capítulo contextualiza a liberdade de expressão na internet com suas potencialidades para a comunicação social, estuda o ambiente de discussão online como um modelo de negócios, abordando as dinâmicas online e as práticas operacionais das empresas que fazem a gestão dos fluxos, insere a democracia, liberdade de expressão, direito a informação e a pluralidade de opiniões nas discussões online e os efeitos das análises algorítmicas no trânsito das informações online diante da modulação e moderação de conteúdo. O terceiro capítulo trata sobre aspectos das fake news como finalidades, temática e motivos pelos quais se proliferam, discorre sobre os modelos de controle dos fluxos comunicacionais e arranjos sociais virtuais como o das milícias digitais e o mecanismo de filtragem algorítmica.

O quarto capítulo correlaciona influências internacionais na construção dos contornos da liberdade de expressão no país, e analisa as formas de enfrentamento à desinformação no Brasil que seguiram da posição de não intervenção estatal para atuação conjunta dos atores públicos e privados, contextualizando o combate às fake news no Poder Judiciário brasileiro. Já o quinto capítulo, relaciona o referencial teórico utilizado com a Constituição Federal de 1988 para analisar a constitucionalidade das filtragens algorítmicas de fake news disseminadas por milícias digitais.

2 LIBERDADE DE EXPRESSÃO E INTERNET

A internet surgiu como um projeto estratégico militar de comunicação descentralizada para que, havendo um ataque nuclear, a comunicação não fosse interrompida tendo em vista a existência de vários pontos de transmissão (CASTELLS, 2020). A conjugação de esforços que tornou possível esse projeto envolveu estudantes, engenheiros, militares, universidades e foi resultado de esforço coletivo de vários setores de inteligência dos EUA, designada como “república dos especialistas em informática” (LOVELUCK, 2018, p. 57)

No meio acadêmico, o intercâmbio entre estudantes e cientistas impulsionou o sonho de tornar o planeta uma aldeia global e assim a liberdade de expressão na internet inaugurava a perspectiva de um plano de realização humana sem intermediadores, um meio de interação acima das fronteiras, dos governos e das leis. Sob inspiração americana do *free speech* e de acordo com o manifesto da Declaração de Independência do Ciberspaço (BARLOW, 1996), as pessoas poderiam obter informação, trocar ideias e definir planos de ação sobre questões planetárias como meio ambiente e fome, num espaço onde todos seriam igualmente considerados e teriam as mesmas oportunidades de fala. Nesse sentido:

Essa internet das origens permaneceu uma referência central, cujos criadores são apresentados como ‘pai fundadores’, estabelecendo-se um paralelo evidente com os arquitetos da Constituição e da nação norte-americana e com os ideais de liberdade individual que lhe são associados. (LOVELUCK, 2018, p. 60)

Pois bem, a gênese libertária da internet em muito foi graças à localização e à origem americanas que compreendiam a liberdade de expressão segundo uma proteção sem paralelo no mundo (ASH, 2017). A interpretação da Primeira Emenda dos EUA como uma garantia absoluta a favor do discurso opera-se porque, segundo a literalidade do texto, não há exceções que justifiquem limitações à liberdade de expressão (SCHAUER, 2016). Some-se a essas circunstâncias, que a construção jurisprudencial americana sobre liberdade de expressão teve início quando o direito internacional ainda era incipiente e por essa razão o entendimento americano sobre a liberdade de expressão se coloca acima de outros valores democráticos como igualdade, dignidade humana e privacidade e, desse modo, a liberdade de expressão americana, como direito absoluto, se constitui numa característica da cultura daquele país (SEDLER, 2016).

A internet trouxe consigo incontáveis benefícios, expandiu o horizonte dos indivíduos e alterou a percepção de espaço e tempo, ao transpor fronteiras geográficas em segundos e permitir o acesso a qualquer tempo ao armazenar informações de forma quase ilimitada e *ad*

infinitum. Nas palavras de Ash (2017, p. 29) a “internet subverte as unidades tradicionais de tempo e de espaço. Ela aproxima o espaço, tornando-os vizinhos virtuais, mas também comprime o tempo. Logo que alguma coisa seja colocada em linha, geralmente fica lá para sempre”. Essas características são ao mesmo tempo fascinantes e perturbadoras, tanto implicam benefícios quanto desafios para a humanidade.

Nesse sentido, a liberdade de expressão foi fortemente impactada tendo em vista a democratização de acesso a partir da versão 2.0. A configuração expandida do ciberespaço⁷ capaz de conectar milhões de pessoas interagindo de forma simultânea e em tempo real ampliou o espaço de divulgação e a velocidade de propagação da liberdade de expressão. Além disso, a memória virtual longeva em que uma manifestação pode ser cessada a qualquer tempo, torna vívida uma postagem como se tivesse sido recém pronunciada, mesmo após longos anos transcorridos desde que foi publicada.

Em virtude da velocidade de transmissão e dos novos horizontes de tempo e espaço da comunicação online, os efeitos decorrentes do exercício abusivo da liberdade de expressão assumem proporções imprevisíveis. Nesse sentido, Dan Wiesel (2020, p. 99) explica que o fluxo de uma postagem online é inicialmente desconhecido, não há uma lógica pré-definida, porque a interatividade entre os usuários e o direcionamento algorítmico reciprocamente se associam na distribuição do conteúdo razão pela qual se desconhece o espectro de divulgação.

Assim, medidas de reparação constitucionalmente previstas no Brasil, com o direito de resposta (art. 5º, V, da CF/88) e as indenizações pecuniárias (art. 5º, X, da CF/88), em regra, não são capazes de suportar os danos porque o caminho de uma postagem em rede é desconhecido e por essa razão é difícil mensurar os danos e aferir os prejuízos sofridos que podem assumir proporções exponenciais, mormente quando a postagem violadora dos direitos permanece acessível e, portanto, continuamente gerando danos.

Uma estrutura baseada uma rede é altamente dinâmica, aberta, capaz de expandir de forma ilimitada, apropriada para a desconstrução e reconstrução contínua da cultura e da política “destinada ao processamento instantâneo de novos valores e humores públicos; e para uma organização social que vise a suplantação do espaço e invalidação do tempo”. (CASTELLS, 2020, p.554)

⁷ “Eu defino o ciberespaço como o espaço de comunicação aberto pela interconexão mundial dos computadores e das memórias dos computadores.” (LEVY, 2018, p. 94).

O ambiente digital é fluido, disruptivo, sempre em movimento, mas a memória é perpétua fazendo com que as postagens se perpetuem de modo indefinido e, enquanto não são removidas, as manifestações lesivas continuam a produzir danos porque podem sempre emergir e repercutir efeitos. O direito de resposta também não se revela eficaz por que, como é difícil rastrear a extensão que a postagem atingiu, a possibilidade de defesa na mesma proporção resta comprometida. Desse modo, as exclusões de conteúdo, mesmo que determinadas judicialmente, dependem de atuação das empresas gestoras dos fluxos comunicacionais que detém os códigos operacionais capazes de interferir no trânsito de informações online.

Na fase inicial de desenvolvimento da internet, até mesmo por desconhecer as repercussões da nova tecnologia, a liberdade de expressão online não teve uma regulação. Desse modo, a ampla liberdade de expressão nos termos da Primeira Emenda dos EUA, implicaram num caos desinformativo com aumento exponencial das fake news.

Na internet informações, opiniões, entretenimento, fatos e ficção circulam juntos sendo difícil discernir a verdade do falso, diferenciar a razão/ciência de opinião e distinguir as informações de fake news, especialmente em razão do aprimoramento das técnicas de produção de desinformação que tornaram a revelação do falso ainda mais laboriosa. Assim, o compartilhamento de fake news entre os usuários foi intenso, fomentando guerras de narrativas.

Como visto, no nascedouro da internet havia a aspiração de uma comunicação livre entre os internautas, de não interferência governamental no trânsito informativo, no entanto, a realidade de um espaço de entendimento se perdeu na pandemia de desinformação e nas discussões sem propósitos que assolam a comunicação online. Ao se manifestar sobre esse quadro, Joel Pinheiro (2019, p. 90) traz: “Chegamos à democracia da informação. Mas por algum motivo, o cenário se assemelha mais a um campo de guerra onde gritamos na cara dos outros e não confiamos em ninguém”.

Timothy Ash (2017) também constata que no ambiente online todos falam e poucos escutam. Pois bem, da forma como se desenvolveu a comunicação social online, não há porosidade no discurso, as pessoas se enfrentam sem ponderar argumentos contrários, despejando as próprias e irredutíveis convicções. Esse desvirtuamento dos propósitos humanísticos da internet ocorreu a partir da aquisição dos *backbones* pelos provedores comerciais na década de 1990 (VALENTE, 2020).

Desde então, a finalidade lucrativa foi incorporada à internet e os sistemas passaram a ser operados, intermediados e geridos por grandes empresas de tecnologia como as atuais

Alphabet, Facebook, Apple, Amazon e Microsoft, conhecidas como “*Big Five*” (MOROZOV, 2020, p. 134). Assim, a expectativa de uma inteligência coletiva libertária degenerou nas práticas corporativas das empresas gestoras dos fluxos comunicacionais.

Superada a fantasia original, paira um ceticismo quanto ao futuro “[a] expectativa moderna de expansão constante das luzes foi substituída por outra, um tanto quanto oposta. Iluminismo obscuro é a expressão que melhor sintetiza a percepção atual do futuro como algo que ameaça o programa humanista” (BERARDI, 2009, p.7). Captando o sentimento de desencanto, Mariana Giorgetti Valente (2020, p. 33) sintetiza que a “perspectiva de uma comunicação completamente desintermediada não se realizou. Todavia, o tipo de intermediário na comunicação digital é bastante diferente do intermediário da mídia de massa, porque o espaço para novas comunicações é, em princípio, ilimitado”.

Nesse sentido, começam a ficar mais evidentes as diferenças entre o ambiente online e a realidade física e, ainda que a tecnologia tenha expandido as possibilidades de interações humanas por meio das redes, isso não significa que a liberdade de expressão esteja mais vigorosa em seu aspecto material (MENEZES, 2020). Nas redes, têm sido constantes as situações em que os limites à liberdade de expressão são postos à prova pela divulgação de material de pornografia infantil, circulação de conteúdo terrorista, fake news, discurso de ódio, vingança amorosa, crimes contra personalidade, invasão de privacidade e *bullying* digital, por exemplo.

Além disso, a adoção de medidas pelas empresas gestoras da comunicação online, como a exclusão de conteúdos e contas, com mecanismos tecnológicos que o Estado não dispõe, tem desafiado os modelos de controle de conteúdo tradicionais e posto em descompasso as medidas constitucionais quanto ao exercício arbitrário da liberdade de expressão.

Se antes da comunicação online a grande ameaça ao poder da palavra ou ao dom do discurso era o Estado, com a internet há uma outra relação tão poderosa quanto, formada entre empresas gestoras dos fluxos comunicacionais e usuários, e regulá-la é um desafio. Além disso, na internet, os mesmos mecanismos que são utilizados com fins democráticos, como a produção descentralizada da informação, por exemplo, podem fomentar a distribuição de fake news por milícias digitais e a atividades criminosas como terrorismo, aprofundando a polarização e radicalização social, servindo à manipulação popular, ao induzimento da opinião pública e a interferência nos pleitos eleitorais.

As fake news se multiplicaram na esfera virtual e, apesar de não ser algo novo na política, a amplitude da divulgação e a velocidade da disseminação da internet potencializam o efeito danoso (MORELLI e POLLICINO, 2020). Compreender o ambiente online sem fantasias ou aspirações libertárias, mas como um modelo de negócios, gerido por empresas privadas, cujas escolhas de arquitetura online são orientadas ao lucro é essencial para analisar o ciberespaço e as repercussões para a liberdade de expressão e a democracia.

2.1 Ambiente de discussão online

As inovações da tecnologia online, desde o acesso através dos grandes computadores aos aparelhos móveis, têm expandido a importância da internet e da comunicação online no cotidiano social a ponto de se afirmar sem receio de exagero que a internet se tornou um elemento onipresente. O modo online passou a ser parte inerente da contemporaneidade, e pode-se mesmo falar em vida “on-life” dado o entrelaçamento dos espaços online e offline (HOFFMANN-RIEM, 2021, p. 25).

Em decorrência da mútua influência entre real e virtual, muitas atividades têm sido incorporadas ao ambiente digital e passaram a ter sua versão tecnológica, como relacionamentos afetivos, trabalho, educação, operações financeiras e até mesmo assistência à saúde. Foram constituídas novas formas de interação como home office, aulas remotas, pagamentos e transferências bancárias online e consultas médicas virtuais, sendo essas atividades reflexos de práticas cotidianas que seguiram a tendência online e passaram a se concretizar segundo a dinâmica virtual.

Ocorre que essa interação entre real e virtual não é uma representação da realidade física-fática na internet, mas uma nova forma de exercer atividades cotidianas, assim uma consulta médica presencial não é o mesmo que uma consulta virtual, o trabalho online também tem outra dinâmica de jornada e avaliação de produtividade, muitas vezes até mais até rigorosa. Do mesmo modo, não há uma equivalência entre esfera pública e ambiente de discussão online.

2.1.1 Modelo de negócios

Nesse processo de migração do real para o virtual, nem tudo se encaixa em conceitos ou experiências conhecidas, no entanto, há uma tendência em tentar compreender as situações e compará-las com algo vivenciado. O ambiente de discussão online não representa a realidade, não há uma correspondência simétrica do real com o virtual, mas apesar disso as duas dimensões se influenciam reciprocamente. Assim, apesar do ambiente online ser um espaço de interação e compartilhamentos não se compatibiliza com a esfera pública habermasiana. Para Habermas a esfera pública é:

[...] um fenômeno social... e não se deixa conceber como instituição nem certamente organização; ela tampouco é uma estrutura de normas com diferenciação de competências e papéis, regulamentação de afiliação de seus membros etc. Muito menos representa um sistema; ela decerto permite fronteiras internas, mas é caracterizada por horizontes abertos, porosos e móveis em relação ao exterior ... Da mesma maneira que o mundo da vida em seu conjunto, também a esfera pública se reproduz mediante ação comunicativa, para a qual é suficiente o domínio de uma linguagem natural... Mas a esfera pública não se especializa nem em um nem em outro aspecto; na medida em que ela se envolve com questões politicamente relevantes, deixa o tratamento especializado aos cuidados do sistema político. A esfera pública se caracteriza antes pela estrutura comunicativa, que se refere a um terceiro aspecto da ação orientada ao entendimento: nem relacionados às funções nem aos conteúdos da comunicação cotidiana, mas ao espaço social produzido pela ação comunicativa.”... “aqueles que agem comunicativamente se encontram em situação que eles ao mesmo tempo constituem com suas interações negociadas de forma cooperativa (HABERMAS, 2020, p. 458-459)

De acordo com os ensinamentos de Habermas (2020), a esfera pública é um fenômeno social (1); que tem a função geral de reproduzir o mundo da vida para formar opinião, sem se especializar ou tomar decisões (2); e se caracteriza por ser um espaço social de estrutura comunicativa orientada ao entendimento (3). Dessas lições, há três aspectos relevantes, o primeiro é que esfera pública está relacionada a um fenômeno social, espontâneo, sem uma sistematização, organização, organogramas ou hierarquias. Em segundo, a ação dialética é cooperativa, em que os participantes assumem um posição de mútua contribuição nos atos de fala cotidianos, sem a obrigatoriedade de serem profundos conhecedores nem lhes ser incumbida a tarefa de tomar decisões que ficam reservadas à esfera política. E por último, refere-se ao espaço social com livre acesso e aberto.

Para Habermas (2020), a finalidade da esfera pública é captar anseios e expectativas latentes na sociedade, apurá-los em opiniões para finalmente ser definida uma posição da opinião pública, nesse sentido, o autor coloca (2020, p. 460) “O que transforma tais opiniões enfeixadas em uma opinião pública é a maneira com que são elaboradas e o amplo assentimento que as sustenta”. Esse processo comunicativo de convergência é inclusivo, por estar aberto a terceiros, e dialogado na medida em que há um assentimento sobre os temas e as contribuições racionais tendem a exaurir as controvérsias.

O ambiente de discussão online ocorre em domínios particulares artificiais, estruturados segundo uma arquitetura algorítmica, orientada a finalidade lucrativa, cuja gestão dos fluxos comunicacionais é resultado hierarquizações algorítmicas que selecionam conteúdo, que os significados chegam prontos, diminuindo a espontaneidade social, e que definem pautas em busca de lucros. Desse modo, o ambiente de discussão online não é um fenômeno social espontâneo como a esfera pública habermasiana, mas projetado por cientistas para fluxos comunicacionais interligados por códigos operacionais matemáticos, que proporcionam um espaço conectado em rede, controlado por empresas privadas.

Também não é orientado ao entendimento. Segundo Vânia Baldi (2018), as práticas online induzem a radicalização das opiniões porque o modelo de negócios lucra com a polarização online. E, de acordo com Francisco Callejón (2022a) como a gestão dos fluxos online está direcionada ao lucro, o conteúdo que gera mais engajamento como curtidas, comentários e likes são potencializados e desse modo fake news e radicalismos são privilegiadas no ambiente online. Ainda segundo autor, a organização e distribuição das informações e opiniões dos usuários, tornam a liberdade de expressão um fator de agressão e radicalização.

Além disso, no ambiente online nem todos têm o acesso livre. O Instagram, por exemplo, tem feito restrições nos termos de usos, no tópico compromisso dos usuários, em que veda a participação na comunidade às pessoas condenadas por crime sexual (INSTAGRAM, 2022). O acesso ao ambiente online exige como pré-condição o aceite a termos e serviços contratuais cujas cláusulas são unilateralmente estabelecidas e que vinculam os usuários (SUZOR, 2018). O aludido autor fez um estudo das cláusulas dos Termos e Serviços de 14 plataformas digitais e, na ocasião, observou que, em alguns contratos, há previsibilidade de exclusão de conta e remoção de conteúdo de modo arbitrário, por critério unilateral. Portanto, o acesso ou permanência do usuário está sujeita a pré-condições dos gestores desses espaços privados e condicionado ao aceite a termos e serviços contratuais pelos usuários, em que as cláusulas são unilateralmente estabelecidas e os vinculam, exigências que não se compatibilizam com o acesso livre e aberto da esfera pública compreendida por Habermas (2020).

De outro modo, Eduardo Magrani e Renan Oliveira (2018) dispõe que o ambiente online é uma esfera pública em potencial, no entanto, os autores constatam que, em virtude do modelo de negócios que faz a gestão da comunicação online mediante análises algorítmicas e utilização de robôs, essas expectativas não se concretizaram, razão pela qual tratam a internet como uma esfera pública digital forjada. Em complemento, Jack Balkin (2020) defende que o ambiente

online é uma esfera pública porque é um espaço em as pessoas expressam e trocam opiniões julgando o que está acontecendo ao redor, mas reconhece que a esfera pública online não tem funcionado bem porque, segundo ele, os valores da liberdade de expressão não estão sendo atendidos.

Na forma como se apresenta, o ambiente de discussão online é parte da concretização de um modelo econômico descrito por Shoshana Zuboff (2020) como capitalismo de vigilância, em que dados pessoais dos usuários são coletados durante a navegação com vistas a formação de perfis para customizar a experiência em rede e oferecer publicidade direcionada. Segundo a autora, os dados coletados online são a matéria-prima de revolução tecnológica enquanto as pessoas são a própria fonte de riquezas. Assim, os recursos a serem explorados são inesgotáveis. Ainda de acordo a autora (ZUBOFF, 2020), o Google foi pioneiro em utilizar os dados de navegação para personalizar experiências na internet e direcionar publicidade, fazendo análises preditivas de comportamento segundo evidências de interesses deixadas como rastros durante a navegação.

A expressão Inteligência Artificial (IA) é ampla, compreendendo desde análises algorítmicas simples a sistemas complexos como aprendizado de máquina. Para Gilberto Almeida (2019, p. 423) a Inteligência Artificial “é estruturada por meio de programas de computador que têm nos algoritmos a origem e síntese de sua operacionalidade (dado que as funções do programa são representadas pelas equações matemáticas que consubstanciam o algoritmo)”. Sistemas que usam IA têm habilidades que ultrapassam a capacidade humana para relacionar e processar dados, analisam grande volume de forma célere, interagem com ambientes dinâmicos, adaptando-se de maneira flexível e aprendendo, coletando dados inclusive de sistemas externos no intuito de realizar uma tarefa específica ou alcançar um objetivo, fazendo comparações, correlações, associações e padrões a partir de fotos, vídeos, textos e dados coletados por celulares e sensores (GUTIERREZ, 2019).

O ambiente de discussão online é um espaço de compartilhamento de mensagens, como textos, imagens e áudios entre os usuários. A questão é que o ambiente de discussão online não é neutro e usa mecanismos tecnológicos tanto para fisgar a atenção no intuito de prolongar a permanência da navegação online tendo em vista que quanto maior o tempo de conexão maior a quantidade dados coletados quanto para direcionar publicidade e conteúdo personalizado de modo a persuadir comportamentos futuros.

As pegadas digitais, que são sinais de cada ação online, são relacionadas e produzem arquivos detalhados sobre cada usuário como sentimentos, pensamentos e interesses

particulares, então, de acordo com análises algorítmicas com o uso de inteligência artificial o conteúdo a ser dirigido ao usuário é personalizado, de modo a captar sua atenção por mais tempo para que fique mais tempo conectado.

Como as análises algorítmicas com o uso de inteligência artificial conseguem apresentar resultados rápidos, mesmo em ambientes dinâmicos, são usados para realizar filtragens de conteúdo online tanto para modular quanto para moderar fluxos. Assim, os gestores dos fluxos comunicacionais online ao organizar e estruturar o fluxo comunicacional de modo personalizado para cada usuário, manipulando opiniões e induzindo comportamentos.

Em breves linhas, o modelo do capitalismo de vigilância descrito por Shoshana Zuboff (2020) oferece serviços online gratuitos e com acesso livre que são custeados pela transferência de dados dos usuários e pela venda de espaços publicitários aos anunciantes. Os dados coletados durante a navegação, são analisados por algorítmicos e o resultado dessas operações direciona conteúdo e publicidade online ao público-alvo, escolhido de acordo com características, interesses e preferências do perfil como idade, localização, padrão financeiro, interesse de compras, pesquisa em buscadores dentre outros. A questão é que esse mesmo mecanismo seleciona e direciona informações que tenham mais afinidades com o perfil, sem oferecer uma visão ampla, e essas práticas tanto reforçam convicções pré-existentes quanto conduzem comportamentos e opiniões de forma subliminar (CALLEJÓN, 2020).

Com a dependência da internet cada vez maior e sendo praticamente todas as transações modernas intermediadas por computadores, o capitalismo de vigilância se tornou um modelo econômico totalizante, sendo a conexão onipresente uma espécie de cárcere que se encaixa na lição de Ash (2017, p. 33) segundo a qual: “[s]e algum Estado ou corporação conhecer tudo o que nós exprimimos a todos, seremos menos livres”. Assim, ao se apropriar toda experiência em rede, os usuários são cada vez menos livres e têm a autonomia individual diminuída e as opiniões maculadas por estratégias tecnológicas de manipulação.

Essa situação ficou evidenciada no escândalo da Cambridge Analytica em que dados pessoais foram utilizados para alterar comportamentos dos eleitores na disputa presidencial americana em 2016. Francisco Callejón (2020) explica que as mensagens subliminares enviadas aos cidadãos democratas, segundo perfis psicológicos previamente definidos, criaram desestímulos ao comparecimento nas urnas tendo em vista que o voto nos EUA é facultativo. Como eleitores democratas compareceram menos às urnas, o candidato adversário Donald Trump que concorria pelo partido republicano foi beneficiado.

Assim, o rumo da comunicação online está distante do que foi idealizado, ao se afastar dos ideais iluministas, se aprofunda num sistema de governança arbitrária e pouco transparente, bem diferente da concepção libertária da liberdade de expressão inspirada na Primeira Emenda (KLONICK, 2017). A influência digital sobre a liberdade de expressão, liberdade de informação e pluralismo dá-se quanto a modulação e moderação de conteúdo que estrutura todo o fluxo em experiências personalizadas. Pesquisa nacional recente sobre a influência dos meios de comunicação aponta que 79% dos brasileiros sempre utilizam o WhatsApp como fonte de informação (DATASENADO, 2019).

Pelo que se observa do resultado acima, a informação, que fundamenta as discussões tanto na esfera pública e no ambiente de discussão online, advém do espaço digital. Assim, apesar da internet ser um ambiente plural, o trânsito de notícias e informações não é livre nem plural como se imagina porque em razão do volume de dados que circula online nem tudo é visível a todos. O sistema de análises preditivas do perfil direciona o fluxo segundo as afinidades de cada um e, desse modo, a diversidade de opiniões que chega aos usuários tende a ser suprimida e sequer tomam conhecimento do que lhes foi omitido (PITRUZELLA, 2018).

Como no ambiente online os usuários tendem a ser alimentados por informações que reforçam suas próprias convicções, criam-se dificuldades à alteridade entre posições divergentes, gerando fragmentação e polarização, inviabilizando o processo de interação dialógica da esfera pública habermesiana. Nas discussões online, as pessoas se dividem e diminuem o nível de tolerância social. As guerras de narrativas, ferramentas para bloqueio de contatos e exclusão por administradores de grupos são características do ambiente online radicalizado em que a liberdade de expressão já não se presta a facilitar a comunicação.

Segundo a teoria do agir comunicativo de Habermas (2019), as ações sociais se dividem em agir comunicativo e agir estratégico. Na esfera pública, as discussões direcionam o agir pelo entendimento, de maneira cooperativa e racional, interagindo de modo a refletir sobre críticas e se dispondo a participar das discussões na medida em que justificam sua participação em prol de opiniões negociadas.

Diferentemente, nas relações virtuais, como se apresentam, as ações sociais na comunicação online se estabelecem segundo de êxito de uns sobre os outros. De acordo com Anderson Schreiber (2020) a força argumentativa do mais eloquente muitas vezes silencia outras vozes, limitando a participação dialogada com outros interlocutores.

Como dito, as discussões online ocorrem num meio em que os fluxos comunicacionais são direcionados por algoritmos. Assim, não é uma atmosfera neutra nem de encontros ocasionais com conversas reflexivas entre pessoas com diferentes pontos de vista que debatem e são permeáveis a argumentos contrários como ocorre numa esfera pública democrática. Além disso, no ambiente online nem todos têm as mesmas oportunidades de serem ouvidos e não há uma visão ampla do tema.

Timothy Ash (2017, p. 68) sintetiza bem a diferença entre os espaços de interação na esfera pública e no ambiente online nos seguintes termos: “A esfera pública possibilitada pela internet é o oposto do ideal liberal segundo o qual os indivíduos são confrontados com fatos inconvenientes, argumentos contrários e valores diferentes(...) forçados a questionar e refinar as próprias convicções.” Segundo Ash(2017), na internet, os usuários têm suas convicções reforçadas por informações se que alinham ao seu ponto de vista sem interações entre posições contrárias. Cass Sunstein (2010) explica que esse direcionamento de informações acaba por polarizar as opiniões e radicalizar as pessoas à medida que se fecham a interações divergentes, causando um distanciando dos que se opõem.

Para Ingo Sarlet e Ivar Hartmann (2019), a comunicação online está sedimentada numa relação de poder, há uma assimetria entre os usuários e as grandes empresas de tecnologia que fazem a gestão dos fluxos comunicacionais em que os últimos manipulam toda a esfera de discussão, seja pela modulação de informações ou pela moderação de conteúdo, impulsionando ou sombreando postagens, procedendo exclusões de conteúdo, bloqueios ou remoções de perfis através de mecanismos tecnológicos de interferência na dinâmica da comunicação online.

Mesmo as relações entre os usuários apesar de serem descentralizadas, multidirecionadas conduzem a uma noção errônea de paridade entre eles porque nos domínios privados dessas empresas nem todos têm as mesmas condições de visibilidade já que os algoritmos promovem algumas postagens e fazem com que sejam mais repercutidas ao passo que sombreiam outras do campo de discussão e, assim, mesmo que a informação esteja disponível online não significa que seja igualmente distribuída a todos.

Ainda que o acesso ao ambiente de discussão online continue aberto, trazendo a falsa impressão de correspondência entre a esfera pública, não há uma reconstrução da realidade física no ciberespaço. A gestão dos fluxos comunicacionais online se norteia segundo finalidades lucrativas, em que empresas gestoras dos fluxos comunicacionais online, perseguem ganhos e não interesses públicos.

Numa análise sobre a relação entre capital e tecnologia, Shoshana Zuboff (2020, p. 27) diz que: “Não podemos avaliar a atual trajetória da civilização da informação sem deixar evidente que a tecnologia não é e nunca deve ser um fim em si, isolado da economia e da sociedade. Isso significa que a inevitabilidade tecnológica não existe”. A autora defende que o problema não é a invenção da internet ou da inovação da comunicação online, mas como a tecnologia foi conduzida segundo escolhas mercantis e direcionadas estrategicamente para permitir ao capitalismo a apropriação de toda experiência humana.

Como espaços de influência mútua, o ambiente de discussão de online repercute na esfera pública, destruindo a força socializadora da ação comunicativa, seja porque determina a pauta de discussão diminuindo a espontaneidade social ou porque desinforma as pessoas com fake news, por exemplo, afetando a tolerância social de uma convivência conciliadora e prejudicando a convergência na tomada de posição segundo ditames democráticos.

Desse modo, o ambiente online não é natural ou desprezencioso, ao contrário, é criado segundo uma estrutura capitalista de ação e de poder, sob uma lógica de dominação e de submissão das pessoas ao poder privado. Todo trânsito online é direcionado a estimular o tempo de navegação para extrair dados, ofertar publicidade e gerar lucros, diferentemente da esfera pública que visa ao entendimento no intuito de instruir a opinião pública na tomada de posição.

2.1.2 Ciberativismo

Com a internet 2.0, em que o usuário é ao mesmo tempo consumidor e produtor de informações (ABRUSIO; MEDEIROS, 2020), a liberdade de expressão ganhou uma nova dimensão, um amplo portal de acesso se abriu às pessoas marginalizadas da mídia tradicional, basicamente rádio e televisão. Para Pierry Levi (2018), houve uma reapropriação pelos indivíduos de uma potência técnica que lhes conferiu autonomia para livres contatos, desvinculados de instituições burocráticas.

Desse modo, a internet deu voz às pessoas e se apresentou como um meio de efervescente participação com compartilhamento online de mensagens, textos, áudios, fotos e vídeos, estabelecendo a comunicação coletiva, rápida, multidirecional, interativa, recíproca e heterogênea. A cibercultura, caracterizada pela difusão do saber descentralizado, da aprendizagem coletiva por meio de troca de experiências em massa e de processos de

colaboração entre os indivíduos avivou o sonho da ciberdemocracia – a participação política por meio eletrônico. (LEVY, 2018)

Em virtude das ferramentas tecnológicas, o acesso do público à cena política foi inédito. A sensação é que as pessoas falam por si mesmas sem intérpretes, jornalistas, parlamentares ou representantes. Além do poder da fala ter sido amplificado pelas dimensões virtuais sem fronteiras e pela velocidade de propagação online, a não edição das manifestações trouxe legitimidade a voz dos usuários, que passaram a dividir as insatisfações, fazer denúncias, compartilhar sentimentos, trazendo visibilidade ao que antes estava latente, mas abafado, um exemplo bem expressivo dessa realidade foi a eclosão do movimento estudantil no Brasil em junho de 2013 (MAGRANI, 2014).

Além disso, o resultado sobre a aprovação ou não das condutas públicas, auferido pelo número de likes e comentários, como uma espécie de termômetro da popularidade dos políticos e satisfação dos cidadãos a respeito de obras e serviços, tornou a internet uma importante ferramenta para definir a atuação política dos congressistas. Os discursos com ideologia, a retórica dos parlamentos ou a figura exuberante de um candidato cederam a mecanismos tecnológicos de captação de votos. Francisco Callejón (2020) descreve como a publicidade online e a propaganda eleitoral subliminar, de acordo com o perfil psicológico do eleitor, pode influenciar as escolhas políticas.

Nesse sentido, Eduardo Magrani (2014) constata que é inegável a influência entre a realidade física e o ambiente de discussão online, no entanto, afirma que não há procedimentos formais de deliberação que possam viabilizar um modelo de ciberdemocracia plena. Segundo o autor supra, mesmo a participação popular direta por meio dos cliques não tem se mostrado mais eficiente que a democracia representativa, apesar do potencial da internet para novos espaços comunicativos e da rápida organização de movimentos e mobilizações populares. De acordo, Byung-Chul Han (2019) afirma que o aglomerado de ciberativistas do espaço digital não tem um comportamento consistente, um plano de ação capaz de promover mudanças a longo prazo.

Assim, o ambiente de discussão online não se confunde com a esfera pública nem com o regime democrático. Conforme os ensinamentos de Habermas (2020), esfera pública e sistema político são distintos. Segundo o autor, na esfera pública os fluxos comunicativos são captados de modo a condensar opiniões para tomada de posição, no sistema político são tomadas as decisões mediante “procedimentos institucionalizados” e ele complementa (HABERMAS, 2020, p. 460) que opinião pública “não é o agregado de opiniões individuais

interpeladas separadamente e manifestadas em privado”, por isso não pode ser confundida com enquete. Desse modo, consultas ou manifestações populares com respostas binárias, algo como sim ou não, são insuficientes para caracterizar opinião pública ou conduzir uma democracia direta.

O ambiente de discussão online, como se mostra na atualidade, não parece capaz de promover um debate racional capaz de produzir um instrumento legislativo negociado entre segmentos sociais diversos. As discussões online não costumam ser equilibradas, nem livres de irritação, desprezo ou indiferença (NOHARA, 2020). No ambiente online não há procedimentos definidos de como seria o processo de criação legislativa, com trâmite definido, quórum para aprovação e prazos, para internalizar as divergências e obter como resultado um instrumento normativo democrático com efetiva participação popular na elaboração.

Numa atmosfera livre, a esfera pública toma posições e as decisões são reservadas ao sistema político cujos procedimentos são previamente institucionalizados e as deliberações conferem legitimidade às decisões políticas (HABERMAS, 2020). Na esfera pública, a opinião pública, apesar de não decidir, influencia o poder político, porque à medida que os parlamentares são filtros que absorvem demandas populares, fazem proposições legislativas e, com base em um procedimento pré-definido de debates e votações estabelecidos constitucionalmente, aprovam ou não um texto de lei.

Num artigo em que analisa o conceito de esfera pública e que relaciona críticas apontadas por outros autores que contestam a aplicabilidade e as falhas da esfera pública defendida por Habermas, Cristiana Losekann (2014) conclui que o conceito de esfera pública em democracias contemporâneas se mostra em evolução e, por não estar consolidado, merece ser aprofundado. Desse modo, ainda que se defenda o ambiente de discussão online como uma versão tecnológica para o modelo de esfera pública é indispensável que existam os elementos definidos por Habermas como aspecto subjetivo do conceito, que é o ânimo de entendimento, e os aspectos objetivos, relacionados a um fenômeno social e a um espaço aberto, não dominado por relações de poder ou submissão que comprometem a liberdade de interação.

O ambiente de discussão online sendo um modelo de negócio não é natural ou espontâneo porque há uma dinâmica de poder e manipulação, por isso não parece se enquadrar no conceito de esfera pública habermasiana. Na internet, os mesmos mecanismos que são utilizados com fins democráticos como acesso livre, descentralização da produção de conteúdo, pluralidade de fontes podem servir a manipulação com a estruturação da comunicação de modo

direcionado, personalizando a informação para cada usuário, induzindo a opinião pública e interferindo nos pleitos eleitorais.

Isso ocorre especialmente porque a pauta de discussão é influenciada por práticas operacionais online obscuras que inconscientemente introjetam temas no ambiente de discussão online com significados embutidos. Além disso, no ambiente de discussão online não há uma separação entre fato e ficção, verdadeiro e falso, ciência/razão e opinião, então todas as manifestações que circulam online são igualmente consideradas na apuração da vontade popular que se revela deturpada pela desinformação que inunda as redes e pelo descompasso entre a realidade vivenciada e as análises algorítmicas com o uso de IA.

Desde a internet 2.0, em que as pessoas produzem, consomem e compartilham conteúdo, os usuários falam por si e se mostram sedentos por apresentar opinião sem intermediários. Nesse cenário, a representação parlamentar tem sido considerada como um entrave tendo em vista que todos querem falar diretamente e não aceitam o filtro dos políticos, pondo em xeque a democracia representativa (HAN, 2019). No entanto, o debate parlamentar, as conversas entre membros de partidos ideologicamente adversários são essenciais para que os desacordos sejam incorporados aos debates legislativos e as deliberações aprovem proposições construídas (WALDRON, 2004). Como se pode perceber, tanto na esfera pública quanto na esfera política há um propósito de entendimento.

A moderna conexão em rede permanece aberta com exponencial volume de informação disponível, mas perdeu o caráter de emancipação coletiva porque houve o desprestígio da verdade factual em detrimento das versões e opiniões pessoais. Desse modo, para Byung-Chul Han (2019) a interatividade perdeu o altruísmo, a preocupação com o todo. Segundo ele, a racionalidade e os usuários se constituem num aglomerado de indivíduos sem alma e sem espírito, sem formações firmes, sem energia política e incapazes de ação para um programa de futuro (HAN, 2019).

A internet continua com acesso democrático e sem edição sobre as manifestações dos usuários, mas a participação nas redes, o “ativismo de clique” não tem demonstrado força para promover um planejamento a longo prazo (HAN, 2019). Além das discussões online não promoverem um debate racional dialógico, o poder de voz não ocorre de forma igualitária. Em razão do volume de informações que circula online, a modulação de conteúdo prioriza algumas manifestações em detrimento de outras e, portanto, não há uma igual consideração entre todos os usuários, sendo ilusória a sensação de participação cidadã igual que a mídia online fomenta.

2.2 Comunicação social eletrônica (Art. 222 § 3º da CF/88)

O direito à comunicação, segundo Camilo Vannuchi (2018), se caracteriza por ser uma via da mão dupla entre os sujeitos envolvidos, ou seja, não se esgota na transmissão da informação, mas na interação e na capacidade de influir sobre a produção da informação, por isso o direito à comunicação não se confunde com liberdade de expressão ou com direito a informação. Por ter uma percepção mais ampla, o direito a comunicação está mais relacionado à troca, à reciprocidade e ao diálogo e não somente a expressar opiniões ou sentimentos e receber informação.

O autor faz uma interessante correlação entre o direito de comunicação e participação política, “a defesa do direito humano à comunicação no início do século XXI é afeto à defesa de um modelo de participação política que se convencionou chamar de democracia deliberativa, mais permeável às demandas da sociedade” (VANNUCHI, 2018, p. 174). Assim, o direito a comunicação refere-se à democratização dos espaços políticos para que sejam mais acessíveis aos cidadãos e permeáveis aos anseios populares.

A Constituição Federal do Brasil de 1988 prevê uma seção especial dirigida à comunicação social e dispõe a respeito da operação de empresas jornalísticas de radiodifusão além de uma menção especial aos modernos meios de difusão de informação em larga escala denominados de comunicação social eletrônica. A atividade de divulgação de informações numa sociedade de massa, plural e democrática, compreende a produção, distribuição e participação popular nas notícias.

Desse modo, o art. 220 caput, da CF/88, que trata sobre a comunicação social reforça a cláusula geral da liberdade de expressão contida no art. 5º, IV, da CF/88. O art. 222 §3º da CF/88, ao se referir especificadamente à comunicação social eletrônica, enuncia que essa modalidade também deve observar os mesmos princípios da produção e programação das emissoras de rádio e televisão, expressos no art. 221 da CF/88 e que estabelece, dentre outros, a finalidade informativa.

Na comunicação social prevalece o entendimento de não interdição prévia sobre as manifestações. A comunicação social online, modalidade da comunicação social eletrônica, foi

concebida segundo a ampla liberdade de expressão e do acesso livre, no entanto, as consequências da divulgação sem checagem prévia de veracidade, de forma veloz e difusa trouxe graves consequências para a imprensa, a esfera pública, a democracia e a liberdade de expressão e forçaram a necessidade de compreensão sobre a finalidade da liberdade de expressão e a respectiva responsabilização dos gestores da comunicação com a atribuição de deveres frente ao fluxo comunicacional online.

Para Giovanni Pitruzzella (2018) a internet alterou a estrutura das informações e a forma como é produzida e distribuída, causando impactos na esfera pública. Segundo ele, duas mudanças são significativas. A primeira está relacionada a produção da informação de forma descentralizada onde os próprios usuários produzem e consomem as informações online. A segunda mudança está relacionada a essencialidade dos serviços capazes de ordenar e facilitar a conexão entre os usuários.

Desse modo, ao tempo em que houve a descentralização na produção da informação com aumento significativo nos fluxos, ocorreu a concentração dos serviços em poucas empresas que tornam a informação disponível e utilizável pelos usuários. (PITRUZZELLA, 2018) Ocorre que as empresas que concentram e fazem a gestão dos fluxos comunicacionais online promovem uma personalização da experiência online e, em termos de distribuição de informações, isso não é uma vantagem porque não canaliza informação balanceada com argumentos prós e contras, afetando o funcionamento do debate público e a formação da opinião pública que se revela fragmentada, sem uma visão do todo.

Ainda segundo Giovanni Pitruzzella (2018), a orientação dos usuários e o rumo dos debates públicos passou a ser influenciado pelos gestores que fazem uma seleção de informações personalizadas, direcionando conteúdo que tenha mais afinidade com a visão de realidade de cada um, restringindo a diversidade de informações sem que o usuário esteja consciente dessas operações. Para o autor, havia uma relação de interdependência entre o pluralismo democrático e a natureza aberta e competitiva do mercado que se enfraqueceu tendo em vista que, com a internet, o mercado está concentrado em poucas empresas, monopólios da informação, e os fluxos são direcionados de acordo com interesses econômicos, sendo selecionados segundo as preferências do perfil do usuário para gerar mais engajamento e tempo de conexão.

Como consequência dessas mudanças provocadas pela internet, houve a alteração do mercado de informações, que deixou de ser produto de organizações de profissionais com padrões jornalísticos, e passou a ser regido mediante análises algorítmicas com o uso de IA,

com uma perda de confiança em relação a mídia tradicional e aos meios de comunicação onde existem controle de qualidade. Isso ocorre porque aquilo que os usuários da comunicação online vêm em seus feeds personalizados, inundados por fake news, não se compatibiliza com as informações balanceadas e plurais da mídia tradicional.

Como os usuários acreditam que as notícias que lhes são dirigidas online são relevantes, confiáveis e neutras, Giovanni Pitruzzella (2018) afirma que houve um declínio da imprensa. Desse modo, o colapso das barreiras que caracterizavam a indústria de informação tradicional, produzida segundo padrões éticos e profissionais de jornalismo, ocasionou um incremento na produção e distribuição de fake news porque cada usuário passou a ser seu próprio censor. A democratização de acesso à internet fomentou a transparência e maior controle do poder público pelos usuários porque trouxe mais visibilidade às condutas públicas, a prestação de contas e ampliou as fontes de informação governamentais, mas, em contraste, a produção descentralizada implicou em muito conteúdo desinformativo online, como as fake news.

Nesse aspecto, Michiko Kakutani (2018), adverte que o multiculturalismo da internet também deu espaço a argumentos radicais de que história e ciência são construções sociais com verdades imprecisas, fazendo crescer o negacionismo histórico e o sentimento de desconfiança em relação as versões oficiais. Ele também afirmou que: “o pós-modernismo não apenas rejeitou todas as metanarrativas, como também enfatizou a instabilidade da linguagem” (p. 65). Desse modo, houve uma convergência de circunstâncias que reinterpretaram a verdade estabelecida como plástica e relativa dando margem a discursos revisionistas e surrealistas.

Nesse sentido, Eugênio Bucci (2019) defende que há uma “estratégia de interdição dos fatos” no sentido de desestabilizar o governo, de se opor ao poder com a supervalorização dos extremos, não somente deturpando o juízo de valor, mas impedindo o juízo dos fatos de modo racional. Segundo o autor, o propósito de contestar e fragilizar o *establishment* é gerar mais lucro com a circulação de mensagens que arrebatam a atenção das pessoas, pondo em dúvida a credibilidade e confiança nas instituições democráticas.

Desse modo, a ampla liberdade de publicar sem checagem prévia de veracidade criou desordem informacional tendo em vista que ficção, opiniões e fake news circulam juntos com fatos e informações, ou seja, do mesmo modo que a internet amplifica a informação também propaga a desinformação e isso traz à tona o debate quanto as possibilidades de controle de conteúdo online e as repercussões sobre a liberdade de expressão.

Episódios ao redor do mundo têm despertado o debate sobre o contingenciamento de modo rápido e efetivo sobre o conteúdo online tendo em vista a necessidade de uma abordagem tecnológica mais célere do que as formalidades do Poder Judiciário e mais restritiva sobre o conteúdo online do que as limitações previstas no texto constitucional.

Como exemplo, podem ser mencionados o escândalo de *Cambridge Analyticals* em 2016 que maculou o resultado eleitoral nos EUA de modo intencional, em que um dos candidatos à Presidência dos EUA utilizou subterfúgios eletrônicos, dispondo de dados psicológicos dos perfis dos americanos, para manipular o comportamento dos cidadãos e influenciar o resultado eleitoral, o genocídio em Mianmar em que a disseminação de discursos de ódio associado a fake news pelo Facebook eclodiu em violência contra a minoria dos *rohingya* ocasionando a morte de milhares de pessoas, também a transmissão ao vivo pelo Facebook do massacre a uma mesquita na Nova Zelândia em março de 2019, e os efeitos do movimento antivax e das fake news que prejudicaram campanhas de imunização na pandemia do COVID-19 (THE GUARDIAN, 2018).

A Seção 230 da Lei de Decência das Telecomunicações dos EUA traz hipóteses legais em que é possível e mesmo obrigatório para as empresas procederem a remoção de material que viola direitos autorais assim como os relacionados a pedofilia e ao terrorismo (NITRINI, 2021). A moderação de conteúdo online pelos gestores dos fluxos comunicacionais *ex post* se mostrou uma medida necessária em virtude dos abusos da liberdade de expressão dos usuários que promovem guerra de narrativas, atingem direitos da personalidade, causam riscos à democracia e a outros “interesses coletivos juridicamente relevantes” expressão utilizada por Wolfgang Hoffmann-Riem (2021).

Desse modo, resta evidente que há uma abordagem diferenciada entre a comunicação social tradicional, do rádio de televisão, e a comunicação social eletrônica online. As mídias tradicionais promovem um controle de conteúdo prévio, estabelecendo pautas e fazendo escolhas editoriais sobre a temática nas matérias jornalísticas, enquanto que, na internet, a liberdade de expressão sofre restrições pelas empresas gestoras dos fluxos comunicacionais em momento posterior à divulgação, sendo um novo tipo de controle editorial (NITRINI, 2021).

O desenvolvimento tecnológico tem criado ferramentas para proibir a publicação de conteúdo online, um controle *ex ante*, utilizando bancos de dados de imagens previamente catalogadas como impróprias, impedindo a transmissão em tempo real de atentados e execuções por meio de filtros ou mesmo dispondo de avisos em mensagens com bullying dando oportunidade ao emissor de refletir antes que chegue ao receptor (DOUEK, 2022). Essas são

medidas restritivas à liberdade de expressão, prévias à publicação, e têm sido amplamente aceitas.

Esses mecanismos de contingenciamento se amoldam as concepções de Stephen Holmes e Cass Sunstein (2019) de que os direitos são meios de auto-organização social e a fruição individual impõe uma conduta “mutuamente respeitosa entre as pessoas” (p. 127), mais adiante afirmam que “embora o direito à liberdade de expressão possa ser reclamado e usado por indivíduos, ele também é pré-condição para um processo eminentemente social – a saber, o da deliberação democrática” (p. 132). Desse modo, é perceptível que a proteção à liberdade de expressão tem dupla dimensão, a individual e a coletiva.

Na mesma obra, numa análise sobre o caráter público das liberdades privadas, os autores (HOLMES e SUNSTEIN, 2019, p. 188) entendem que os direitos “são instrumentos forjados para servir a interesses humanos” e dessa forma se prestam tanto para a realização individual quanto em benefício da sociedade razão pela qual o exercício da liberdade de expressão não pode ser concebido como se o indivíduo fosse um ser isolado, como que tendo direito ao gozo absoluto ou desvinculado de responsabilidades. Diante disso, a par das interdições à liberdade de expressões assumirem um caráter excetivo, as dimensões da comunicação social online traçam novas perspectivas para limitações acetáveis e mesmo desejáveis.

2.2.1 Informação em larga escala

O Brasil é um país com uma população de mais de 200 milhões de habitantes e pelas dimensões territoriais tem proporções continentais sendo culturalmente diverso. Nesse contexto, as informações precisam atingir um público socialmente desigual e geograficamente disperso. Desse modo, mostra-se imprescindível a comunicação que atinja com profundidade esse público mesmo em locais distantes e a internet adquire dimensões inéditas tanto em alcance, velocidade na transmissão e pluralidade de informações.

A cultura democrática online foi absorvida pelos canais governamentais, a facilidade com que a informação circula no meio virtual e as ferramentas para consultas e enquetes eletrônicas tornam o ambiente digital ideal para captar as expectativas e insatisfações populares, viabilizando mais integração do cidadão à vida política por meio de acessos diretos com órgãos oficiais. No Brasil, a Lei de Acesso à Informação (LAI) (Lei nº 12.527/2011), nos art. 7º e 8º, determinou que órgãos e entidades públicas promovam a divulgação de informações em sítios

oficiais na internet e há uma tendência mundial em quantificar dados e agregá-los em bases informatizadas.

Na zona de influência recíproca entre o mundo dos fatos e o ambiente online, a participação popular por meios eletrônicos tem trazido maior transparência aos gastos e senso de responsabilidade aos gestores em virtude da vigilância onipresente que os meios tecnológicos propiciam. Além disso, têm influenciado os gastos orçamentários e as políticas públicas, a título exemplificativo é possível mencionar a consulta pública no período de 24 de dezembro de 2021 a 02 de janeiro de 2022, realizada no site do Ministério da Saúde do Brasil, sobre a vacinação de crianças entre 5 a 11 anos contra a COVID-19 no país (BRASIL, 2021a), em que a população se manifestou favorável mesmo diante da desinformação que circulava online sobre os efeitos colaterais das vacinas e da posição contrária do chefe do Poder Executivo.

Há um crescente engajamento popular através da internet às decisões governamentais, demonstrando o potencial do ambiente online para canalizar as expectativas e insatisfações sociais, viabilizando mais integração do cidadão à vida política por meio de acessos diretos aos órgãos oficiais. A comunicação do público com sites oficiais e redes sociais⁸ dos parlamentares pode restaurar a credibilidade política à medida em que os representantes se mostram mais acessíveis aos eleitores e sensíveis às suas demandas.

Desse modo, a comunicação da informação através da internet atende aos critérios de interação social e variedade cultural, tendo em vista que os usuários ativamente participam da produção das informações e através de compartilhamento e likes interagem e amplificam o alcance. Em regra, o acesso à internet é democrático, embora haja uma seleção posterior em que as exclusões de postagens e contas são cada vez mais comuns.

É inegável que numa sociedade de massa e plural a divulgação da informação depende de meios tecnológicos e, especialmente depois da democratização do acesso, a internet se mostrou um meio importante para dar publicidade e transparência das condutas públicas. Além disso, a multiplicidade de serviços online e a pluralidade de fontes de informação facilitou o controle das contas públicas, a pesquisa de dados, a participação em audiências públicas,

⁸ Utiliza-se o conceito legal neste trabalho: “rede social: aplicação de internet que realiza a conexão entre si de usuários permitindo a comunicação, o compartilhamento e a disseminação de conteúdo em um mesmo sistema de informação, através de contas conectadas ou acessíveis entre si de forma articulada.” (Art 4º, X, do PL 2630/2020)

acompanhamento de processos informatizados, comprovação de situação eleitoral online dentre outros serviços.

Ocorre que apesar do potencial da internet para permitir o acesso à informação de fontes variadas, inclusive oficiais, e ainda, tendo viabilizado o intercâmbio de modo direto, rápido, síncrono e sem formalidades burocráticas, mesmo que entre os usuários remotos, é preciso cautela quanto ao julgamento dos serviços e das informações disponíveis online porque há quem se valha dessa aparente imparcialidade e neutralidade para disseminar desinformação nesse ambiente. Nem todas as informações que circulam online são verificadas, verdadeiras ou confiáveis porque os padrões jornalísticos e as barreiras editoriais foram dispensadas na comunicação online.

Com a internet 2.0, qualquer indivíduo passou a ser um potencial produtor de informações, assumindo um papel ativo na comunicação online, sendo cada um seu próprio censor, e isso acarretou aumento exponencial da quantidade de fake news, cujo conteúdo é desinformativo. As operações algorítmicas não apenas priorizam conteúdo desinformativo nos fluxos comunicacionais, mas também direcionam informações ao feed de notícias dos usuários de acordo com as preferências do perfil, diminuindo a pluralidade de fontes e fragmentando informações, oferecendo apenas aquilo que adere as opiniões e concepções previamente catalogadas.

Além disso, Timothy Ash (2017) argumenta que, apesar do acesso aberto, nem todos tem as mesmas oportunidades de serem ouvidos, por isso Owen Fiss (2005. p. 49) defende que “algumas vezes nós devemos reduzir o volume de alguns para podermos ouvir vozes de outros.” Assim, apesar da manifestação online ser facilmente publicada pelo usuário, o caráter desigual da distribuição dos fluxos em redes não assegura que a postagem atinja o público almejado, prejudicando o direito de informar.

Segundo Balkin (2018), apesar da descentralização da produção de conteúdo ter aumentado exponencialmente o trânsito de informações online, toda a comunicação é concentrada em poucas empresas cuja gestão dos fluxos é autoritária, no entanto, por disporem de tecnologia para modular e moderar o conteúdo online se tornaram essenciais. O autor explica que essas empresas organizam as conversas públicas e facilitam a participação dos cidadãos online bem como estruturam a informação, dispondo quais assuntos e em que ordem serão apresentados no feed de notícias.

Giovanni Petruzzelli (2018) explica que, para tornar a informação acessível e utilizável, as empresas que fazem a gestão dos fluxos comunicacionais desenvolveram algoritmos que selecionam conteúdo de acordo com dados do perfil facilitando as buscas, reduzindo o tempo e os custos de pesquisa, tornando as compras mais fáceis e seguras além de oferecer maior satisfação nos intercâmbios com outros usuários. Ainda segundo o autor, as operações algorítmicas filtram as notícias e opiniões do mundo exterior construindo realidades diferentes, individualizando-as para cada usuário, de acordo com cada visão particular. Como consequência, a diversidade e a pluralidade também são filtradas e cada usuário-cidadão só é exposto ao que está em sintonia com suas concepções de modo que suas crenças e opiniões são fortalecidas.

Diante dessas escolhas sobre os fluxos online, Loveluck (2018) afirma que a ilusão de uma internet neutra não existe mais. Segundo o autor, desde que a banda larga congestionou o tráfego de dados, especialmente em razão dos vídeos que aumentaram o volume de dados em circulação, foi necessário estabelecer diferentes níveis de prioridade sobre o fluxo, determinando por filtros o que terá mais visibilidade. Desse modo, para ele, a interferência sobre o trânsito de dados, que passou a ser selecionado, fez ruir também a percepção original de que, por ser neutra, a internet era sinônimo de liberdade, igualdade e justiça. A forma como a informação online é distribuída não é neutra ou imparcial.

Desse modo, o tratamento online da informação como uma mercadoria ajustável a vários públicos traz o risco de tornar a informação de qualidade, verificada ou apurada segundo padrões profissionais de jornalismo numa mercadoria de luxo, causando disparidades para o ambiente democrático, de modo que, quem puder pagar terá acesso a melhor conteúdo. Nesse sentido, informação de qualidade, verificada, não pode ser artigo de raro ou de difícil acesso, a ser peneirado diante do mar de desinformação que assola o ambiente digital.

Em regimes democráticos, as informações devem ser dirigidas sem distinções dos grupos populacionais, ou ao menos com a advertência de que o fluxo não é uniformemente distribuído para todos, de modo que torne claro que as informações são fragmentadas, tornando transparentes os critérios que motivam os direcionamentos e dando oportunidade para os usuários alterarem dados do perfil no intuito de democratizar o que lhes é enviado.

Ao discorrer sobre o futuro das redes, Emanuel Castells (2020, p. 498) defende que “A arquitetura da rede é, e continuará sendo aberta do ponto de vista tecnológico, possibilitando amplo acesso público e limitando seriamente restrições governamentais ou comerciais a esse acesso, embora a desigualdade social se manifeste poderosa no domínio eletrônico.”, ou seja, o

acesso permanecerá aberto mas a velocidade de navegação, a distribuição das informações desbalanceadas estratificarão os usuários/cidadãos segundo desigualdades econômicas, quem puder pagar terá acesso mais rápido e informações melhores.

Informações selecionadas para grupos populacionais segmentados criam ambientes fragmentados e, assim, algoritmos fabricam artificialmente mundos particulares onde não deveria haver divisões, destruindo a realidade socialmente compartilhada. O problema é que, por não ser transparente o trânsito online, a maioria dos usuários desconhece como ocorre a gestão dos fluxos e acreditam que as informações que lhes são apresentadas são confiáveis, relevantes e neutras, não tendo consciência de que lhes é omitido e de como o ambiente online é fragmentado (PITRUZZELLA, 2018).

A democracia requer diversidade e internaliza o conflito no processo de tomada de decisão, desse modo a pluralidade de fontes é essencial para que sejam retratadas as condições sociais e garantir a legitimidade do consenso. No entanto, por não ser uma representação fiel do mundo real, o ambiente de discussão online ainda que tenha o potencial para divulgar informações plurais nem sempre as dirige ao usuário e, assim, as pessoas ficam presas às suas concepções originais, reforçando ideias e sentimentos já existentes sem promover a alteridade de pontos de vista, ou seja, ainda que o intercâmbio online seja intenso, não há efetiva troca nas interações.

Portanto, o direcionamento algorítmico de informações online desvirtua o direito de informar e de ser informado na relação democracia-cidadania. Isso ocorre porque as informações direcionadas aos usuários são fragmentadas, reproduzindo preconceitos, e nem todos os usuários têm a mesma visibilidade porque o alcance de uma manifestação é manejado buscando lucros, sem que sejam distribuídas igualmente. Assim, a finalidade político-democrática da informação é obstaculizada e não atinge o fim a que se propõe que é o de contribuir para o debate entre posições divergentes e convergentes no intuito de construir decisões conciliadas.

As informações estão disponíveis na internet, inclusive oriundas de fontes oficiais e verificadas, mas como o volume de dados é superior a capacidade humana de processamento, há uma seleção operada por análises algorítmicas com o uso de IA que determina o que será direcionado a cada usuário. A governança online é operada por empresas com finalidade lucrativa e, por não se guiar por práticas democráticas, a distribuição da informação não é neutra ou livre, mas condicionada a interesses econômicos que direcionam fluxos de acordo com perfil do usuário.

Além disso o ambiente online tem muita desinformação já que não existem barreiras editoriais prévias. Há um contraste, ainda que a internet seja potencialmente propícia para divulgação de informações em larga escala é congestionada por muita desinformação e os fluxos priorizam conteúdo que provoca agitação como fake news e que se adequam ao perfil do usuário ao invés de informações balanceadas.

2.2.2 Algoritmos e grande volume de dados

A comunicação online é uma ruptura sobre os modos até então conhecidos de comunicação e se difere das mídias tradicionais. A televisão e o rádio são veículos de massa em que há direcionamento unilateral da informação, a escolha de pauta é feita por seleção editorial segundo padrões jornalísticos. A internet, com seus avanços tecnológicos desde a versão 2.0, tornou a comunicação interpessoal síncrona e horizontal, sem interlocutor visível para dar significado ao conteúdo das mensagens e com múltiplas interações.

A comunicação online é gerida por empresas privadas que fazem a gestão dos fluxos comunicacionais perseguindo lucros, assim utiliza-se análises algorítmicas com o uso de IA, que não são visíveis, para administrar o fluxo de informações online e direcionar conteúdo. O modelo de negócios consiste na venda de publicidade e na extração e apropriação de dados dos usuários, então quanto mais tempo o usuário permanecer online maiores serão as quantidades de dados de navegação coletados e mais eficiente será a publicidade direcionada segundo interesses do perfil do usuário.

Para tanto, os sistemas de análises de dados com o uso de IA procedem leituras de grande volume de dados de forma rápida, apresentando soluções com eficiência matemática para os objetivos propostos mediante o uso de inteligência artificial. Ocorre que esses sistemas autônomos inteligentes procedem leituras dos dados apresentando soluções, gráficos e tomando decisões segundo análises obscuras à inteligência humana e, desse modo, interferem em comportamentos, condicionam escolhas e selecionam informações prejudicando juízos de valor, manipulando a opinião popular e desvirtuando decisões públicas com técnicas subliminares de indução e predição de comportamentos.

Em virtude da opacidade das decisões geradas por inteligência artificial, Nilton Correia da Silva (2019, p.47) aponta que as pesquisas atuais têm direcionado não apenas à assertividade

(=eficiência) mas também à interpretabilidade (=compreensão) das soluções artificiais de modo que possam ser auditadas por um observador humano.

Para Juarez Freitas e Thomas Freitas (2021, p. 59), a inteligência artificial deve ser usada para o bem e em defesa do humano, desse modo a regulação tecnológica deve ser direcionada para tornar as relações humanas menos artificiais e obscuras, para que a “aprendizagem (relativamente) autônoma da máquina catalise a inclusão informacional e desvele, em tempo hábil, os passos lógicos das decisões algorítmicas de relevância sistêmica.”

Os autores (FREITAS e FREITAS, 2021) aduzem também que, no atual estágio de desenvolvimento da inteligência artificial, as máquinas dependem de uma atividade humana prévia que as calibram. Reforçam que há uma atuação e decisão humana anterior às programações dos sistemas e às soluções matemáticas, argumentam que os parâmetros iniciais definidos por atividade humana condicionam e interferem nas respostas e por isso defendem que princípios éticos sejam observados também no desenvolvimento das tecnologias.

Eles também alertam que é preciso desbaratar as forças tecnológicas e econômicas por detrás das soluções matemáticas que impõem normas técnicas de modo a se sobreporem ao poder estatal, nos seguintes termos: “Nesse quadro de acentuada desmaterialização, justifica-se a regulação que protege, vigorosamente a liberdade contra qualquer subjugação virtual, por ação ou omissão, nacional ou transnacional” (FREITAS E FREITAS, 2021, p. 59).

A análise algorítmica é uma realidade da modernidade e tende a expandir os usos e aplicações em razão do avanço da digitalização que produz grande quantidade de informações e que demanda alto poder de processamento, necessitando de respostas rápidas mesmo diante de volumosos bancos de dados. Por essas razões o uso da inteligência artificial tende a crescer e ser aplicado em escala por conta do volume maciço de dados continuamente produzidos e que dependem de ferramentas cognitivas artificiais para análises e correlações. (GUTIERREZ, 2019).

O aperfeiçoamento tecnológico tem conduzido a coleta e cruzamento de dados entre sistemas, nesse sentido a Big Data amplia o espectro da extração de dados e procede leituras de dados capturados em aplicativos móveis de mídias sociais e em bases de dados privadas, alcançando dados em sistemas online externos (STEIBEL; VICENTE E JESUS, 2019, p.54/55). Nos termos do art. 21 da Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seus Ambientes, a Big Data executa o “tratamento de quantidades consideráveis de informação, organizada em bases de dados.”

Para Emanuel Castells (2020) os dados são a matéria-prima da revolução tecnológica do Século XXI assim como as fontes de energia, carvão e petróleo foram das revoluções industriais anteriores. Nesse sentido, Wolfgang Hoffmann-Riem (2021) constata que os dados digitais não têm significado, a princípio, mas como contém “informações codificadas” se tornam o bem econômico na sociedade de informação.

Desse modo, reproduzindo a metáfora contida no art. 19 da Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seus Ambientes, em que os dados são comparados ao petróleo, Hoffmann-Riem (2021, p. 19-21) lista as vantagens dos dados digitais e conclui que os usos e possibilidades são maiores do que os do petróleo.

Para o autor, a primeira vantagem é que os dados não são finitos porque estão em constante produção, são fáceis de serem extraídos e o valor deles só aumenta. Em segundo, o processamento de dados transmite conhecimento e gera mais poder à medida que podem ser continuamente extraídos. Em terceiro, os dados estão armazenados em bolsões como empresas de tecnologia e governos assim como em computadores individuais, portanto mais acessíveis do que os recursos naturais a serem extraídos das profundezas. Em quarto, podem sofrer vários refinamentos tendo em vista que a cada leitura os dados podem gerar novos produtos. Em quinto, os dados coletados não implicam em consumo ou refinamento imediato. Em sexto, podem ser mantidos para processamento posterior. São, portanto, vantagens significativas.

Nesse sentido Castells (2020), esclarece que o diferencial da atual revolução são as tecnologias de processamento das informações, de modo que o conhecimento acumulado é capaz de gerar novos conhecimentos e informações num ciclo de realimentação cumulativo entre inovação e uso. Há desse modo um processo infinito de geração de informações pelo uso e reuso dos dados, ainda que sejam os mesmos, em sucessivas leituras e análises.

O que caracteriza a atual revolução tecnológica não é a centralidade do conhecimento e da informação, mas a aplicação desses conhecimentos e dessa informação para a geração de conhecimentos e de dispositivos de processamento/comunicação da informação, em um ciclo de realimentação cumulativo entre inovação e seu uso. (CASTELLS, 2020, p. 88)

Assim, a tecnologia gera conhecimento de modo exponencial porque, a partir de dados arquivados, é capaz de traçar perfis dos indivíduos e atualizá-los constantemente, prever comportamentos segundo interesses recentes e direcionar conteúdo de acordo as predileções pessoais mais atuais. Através de leituras sucessivas, há um contínuo processo de renovação dos perfis e direcionamento de conteúdo.

No livro intitulado *Algoritmos de Destruição em Massa*, Cath O’Neil (2020), argumenta que as aplicações matemáticas conduziam a economia de dados com a expectativa de ganho de tempo, imparcialidade e objetividade nas análises, no entanto, ela revelou que as combinações algorítmicas da inteligência artificial, por imularem o raciocínio humano, embutiam vieses, preconceitos e equívocos, ainda que sem a intenção discriminatória ou mesmo em busca de objetivos louváveis.

Para O’Neil (2020, p. 47), “os algoritmos são, por projeto, caixas pretas impenetráveis”, mas, no entanto, e apesar disso, são usados em várias atividades cotidianas como análises curriculares, assistentes de diagnóstico por imagem, sistemas de transações financeiras, dentre outras. Assim, acesso a sistema público de saúde, preenchimento de vaga de emprego e concessões de empréstimos bancários, por exemplo, podem ser determinadas segundo critérios não transparentes de julgamento e isso compromete a contestação dos resultados.

As empresas que desenvolvem a inteligência artificial por outro lado resistem à abertura dos códigos fonte para auditabilidade de suas práticas e alegam que se trata de sigilo comercial, já que investiram no desenvolvimento dos sistemas e detém a propriedade intelectual de suas invenções. Defendem ainda que, revelado o código-fonte, a segurança dos clientes resta comprometida porque sujeita ao vazamento de dados e a divulgação não autorizada de informações pessoais como senhas, movimentação financeira, sigilo das comunicações telefônicas dentre outras.

Como referido acima, sistemas automatizados têm definido muitos aspectos da vida em sociedade e por consequência muitos comportamentos têm sido determinados por normas técnicas que acabam por estabelecer normas gerais com poder equivalente ao de lei. As normas técnicas não são elaboradas por instâncias políticas tradicionais, e mesmo assim fixam as práticas e prescrições em matérias como saúde, emprego e crédito (FRYDMAN, 2018, p. 59). Diferentemente, as regras jurídicas passam pelo crivo de legitimidade democrática, uma vez que são deliberadas por órgãos de representação popular.

Ao tratar sobre as normas técnicas, Benoit Frydman (2018) desmistifica e critica algumas características que lhes são comumente atribuídas. Primeiro ele afirma que as normas técnicas não regulam apenas coisas ou objetos, também dirigem homens e seus comportamentos. Depois esclarece que a obrigação de as respeitar não é voluntária e que, prática, impõem a mesma força constrangedora da obrigação jurídica. Por último, argumenta que padrões e indicadores técnicos não são neutros nem democráticos, haja vista que muitos são elaborados de maneira unilateral por uma única empresa ou clube e, em virtude da alegada

natureza qualificada, escapam ao controle judicial, faltando-lhes, assim, legitimidade. Ele ainda critica a posição subordinada e subsidiária que aparentemente as normas técnicas ocupam em relação ao direito e alega que é preciso redefinir esse sistema.

As normas técnicas costumam ser compreensíveis apenas por profissionais especializados, mas, como significam escolhas políticas e econômicas, acabam por determinar comportamentos, e justamente por isso é preciso definir critérios para que a inteligência artificial seja manejada de modo transparente e a evitar decisões artificiais juridicamente condenáveis.

No tocante a liberdade de expressão, para impedir que seja tolhida injustificadamente, é mister que se estabeleçam as situações em que são admissíveis restrições em virtude de colisões com outros direitos fundamentais, diante disso é preciso especificar os parâmetros de exclusões de postagens para que, na interação com ambientes dinâmicos como as redes, as variações das aplicações algorítmicas não incorram em decisões carregadas erros e vícios nos resultados.

No Brasil, a Resolução nº 332 do CNJ de agosto de 2020 trata sobre a aplicação da inteligência artificial no Poder Judiciário e no sentido de contribuir com a “agilidade e coerência do processo de tomada de decisão”, além de reforçar a necessidade de compatibilidade com os direitos humanos (art 4º). Diante da tecnologia, defende “critérios éticos de transparência, previsibilidade, auditoria” (art 8º), imparcialidade, justiça substancial e condições para eliminar ou minimizar erros de julgamento decorrentes de preconceitos (art. 5º).

O artigo 7º, caput, da Resolução nº 332 do CNJ enuncia que decisões judiciais devem preservar a igualdade, a não discriminação, a pluralidade e a solidariedade, e os §§ 2º e 3º do mesmo artigo referem-se às medidas corretivas quanto aos vieses discriminatórios identificados discorrendo sobre a descontinuidade do modelo de inteligência artificial se não houver meios para eliminação dos vieses discriminatórios.

Outros artigos da mesma resolução reforçam a necessidade de segurança dos dados e fontes seguras, estabelecendo as fontes governamentais como preferenciais (art 13), bem como o direito à revisão da decisão (art. 18) e à explicabilidade da decisão artificial (art. 19). Interessante o teor do artigo 23 em que expressamente declara que a utilização de inteligência artificial em matéria penal não deve ser estimulada sobretudo em decisões preditivas. Aludida resolução traz ainda conceitos técnicos como o de usuário, inteligência artificial e algoritmo.

Essa resolução reconhece a existência de várias “iniciativas que envolvem Inteligência Artificial no âmbito do Poder Judiciário e a necessidade de observância de parâmetros de governança, desenvolvimento e uso éticos”. Pode-se dizer que inteligência artificial é uma ferramenta cotidiana e mesmo instituições solenes e formais como o Poder Judiciário não dispensam a assertividade e celeridade de seus modelos, inclusive, em atividades decisórias. Tribunais Superiores como STF e STJ já dispõem de sistemas inteligentes para a triagem dos recursos, como o Victor e ATHOS, que identificam temas de repercussão geral no caso de recurso extraordinário (OLIVEIRA; SOUZA, 2019) e agilizam a formação de precedentes nos recursos especiais (NORONHA, 2020).

A Lei Geral de Proteção de Dados (LGPD), no artigo 20, também faz referência à inteligência artificial ao garantir o direito a explicabilidade das decisões tomadas unicamente com base no tratamento automatizado de dados pessoais. Essa diretriz é intimamente relacionada à transparência das decisões algorítmicas e à extensão do dever de fundamentação e motivação das decisões, tornando-o obrigatório nas decisões artificiais. O direito à explicação se apresenta como condição para outros direitos tais como o da recorribilidade e da reversibilidade decisória porque a contestação com potencial para alterar a decisão só é possível se for capaz de contradizer os motivos sob os quais se apoiou.

Outro ponto interessante levantado dessa vez por Juarez Freitas e Thomas Freitas (2021) no tocante ao direito à explicabilidade trazido pela LGPD refere-se ao reconhecimento de que a “IA pratica atos jurídicos” equiparando a decisão algorítmica “a ato administrativo pra fins de impugnação” e assimilando a “decisão algorítmica como atividade iminente jurídica, não meramente geradora de fatos jurídicos” e, portanto, sujeita, “por analogia, à observância dos requisitos de validade do negócio jurídico” previstos no art. 104 do Código Civil. (FREITAS; FREITAS, 2021, p. 105-107)

No âmbito internacional, a Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais trouxe importantes diretrizes sobre a temática. Existe uma preocupação com injustiças nas decisões em que se utiliza a inteligência artificial, perpetradas por discriminações motivadas por fatores relacionados a raça, localização, condição econômica, religião, preferência sexual, atividade sindical dentre outros. O documento aborda a necessidade de medidas corretivas e não proíbe o uso da inteligência artificial, ao contrário, expressamente incentiva a aprendizagem automática e análises científicas multidisciplinares para combate das distorções.

A Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seu ambiente, data de dezembro de 2018, relaciona 05 princípios: respeito aos direitos humanos; da não discriminação; da qualidade e segurança; da transparência, imparcialidade e equidade; e, por último, o do “sob controle do usuário”. Desses princípios decorrem outros como direito ao julgamento justo, do contraditório, da motivação das decisões dentre outros.

Além disso, avalia situações em que a inteligência artificial já é aplicada e sugere usos, como motores de busca para a melhoria da jurisprudência, opções de pesquisa que relacionem julgados com várias fontes legislativas, criação de chatbots para acesso à informação sem substituição humana, elaboração de indicadores de desempenho e melhoria dos serviços com a adequação dos insumos às necessidades reais.

No Anexo II da referida Carta, dentre as utilizações possíveis da inteligência artificial que exigem precauções metodológicas consideráveis, trata sobre a resolução de litígio em linha, afirma que é preciso deixar claro se o tratamento do litígio é totalmente automatizado ou há participação de mediador ou árbitro, assim como fazer sempre a menção que não se trata de um tribunal, mas de um serviço de resolução alternativa de conflitos. No anexo III traz um glossário sobre os conceitos tecnológicos mais recorrentes.

Também no Anexo II, no tocante ao uso de algoritmos no âmbito penal, o documento considera viável o tratamento de dados para prevenção da criminalidade assim como para estabelecer o tipo da pena alternativa segundo o perfil do indivíduo. No entanto, alerta para o uso com as mais extremas reservas, em matéria penal, nas hipóteses em que os algoritmos avaliam os riscos de reincidência do indivíduo.

No Apêndice I, nos arts. 5, 6 e 7, ao tratar sobre a justiça “preditiva” que, cujo fenômeno, em verdade, está relacionado a previsibilidade de sucesso ou fracasso de um caso sob judge, aferida por modelação estatística de decisões anteriores, menciona o tratamento de dados de jurisprudência de massa e os relaciona aos benefícios e aos riscos onde se contrapõem transparência e previsibilidade versus limitações e parcialidades do raciocínio do software. Além disso, aborda o efeito de possível padronização das decisões em que a fundamentação do caso concreto pode ser desprestigiada em detrimento da uniformização dos julgados.

Como se percebe, a tecnologia evolui a passos largos e o potencial de suas utilizações ainda é desconhecido, no entanto, todo o fascínio que desperta não pode ser romantizado como se fosse apenas em benefícios à humanidade. A descentralização da produção de informação online, o processo de digitalização em escala global e a quantificação da atividade humana em

dados tem incrementado o volume de conteúdo online e por essa razão as análises algorítmicas, com o uso de inteligência artificial, se tornaram indispensáveis para processamento desses fluxos em resultados concretos. Ainda que muitas vezes essas análises algorítmicas com o uso de IA procedem associações e combinações incompreensíveis à cognição humana, e não sejam visíveis ou perceptíveis, as soluções apontadas têm se caracterizado como normas de comandos gerais que moldam comportamentos e decidem muitos aspectos cotidianos.

2.3 Moderação de conteúdo, *lato sensu*, no ciberespaço e o desafio de fazer respeitar direitos fundamentais no ambiente virtual

O ambiente online opera segundo regras próprias cuja dinâmica de interação é determinada por algoritmos de modo que as informações dirigidas ao perfil constituem uma experiência customizada para cada usuário. De igual modo a rede de contatos também é monitorada e os encontros virtuais não são propriamente espontâneos ou ocasionais. A modulação do conteúdo é a atividade de personalização do que cada usuário vê em tela, filtragens direcionam conteúdo tendo em vista o grande volume de informação que circula online, desse modo o teor das informações e contatos dos feeds são selecionados de acordo com as predileções de cada perfil.

As empresas que fazem a gestão dos fluxos comunicacionais online atuam também moderando conteúdo e perfis, procedendo a exclusões e bloqueios de contas e de conteúdo que violam suas políticas internas agindo de maneira proativa ou mediante denúncias de usuários, além de dimensionarem o fluxo online conferindo maior ou menor visibilidade do conteúdo. Evelyn Douek(2022) trata como métrica de prevalência, o dimensionamento da amplitude de divulgação, impulsionando ou sombreando conteúdo.

Os autores Terry Flew, Fiona Martin e Nicolas Suzor (2017, p. 40) traçam um conceito mais abrangente de moderação de conteúdo, que engloba tanto a modulação quanto a moderação de conteúdo, nos seguintes termos: “a triagem, avaliação, categorização, aprovação ou remoção/ocultação de conteúdo online de acordo com as políticas de comunicação e publicação relevantes. Procura apoiar e reforçar o comportamento de comunicação positiva online e minimizar a agressão e o comportamento antissocial.”

Nesse trabalho, as atividades de modulação e moderação de conteúdo foram tratadas de modo diferenciado. A modulação com a seleção e direcionamento de conteúdo. A moderação

como a exclusão, remoção, bloqueio, impulsionamento e sombreamento de conteúdo. Convém reforçar para o leitor que a abordagem mais abrangente e mais usual pelo referencial teórico considera as atividades de modulação e moderação como sinônimas, referidas sob o mesmo termo moderação. Diante disso, quando forem referidos como atividades conexas, moderação e modulação, serão tratada na dissertação como moderação de conteúdo *lato sensu*.

A par dessas conceituações, percebe-se que o ambiente virtual não é tão livre como parece, a utópica neutralidade da rede de não intervenção no fluxo comunicacional se mostra cada vez mais distante da realidade pragmática tendo em vista que as interferências se expandem em ritmo acelerado. Bem verdade que, desde a internet 2.0 em que os usuários produzem e consomem informações não há edição sobre as postagens, no entanto, o trânsito de informações tem sido afetado para corresponder a ganhos econômicos.

2.3.1 Expansão dos critérios de modo discricionário sem uma diretiva regulatória específica

O ambiente online não é natural, sendo a configuração atual uma arquitetura em rede projetada para capturar dados online e apta a controlar e vigiar toda experiência de navegação. Como os gestores dos fluxos online são empresas privadas com finalidades lucrativas, os algoritmos são orientados a perseguir ganhos, seja direcionando publicidade ou se apropriando de dados dos usuários.

Em um artigo, Kate Klonick (2017) revela que nos EUA há uma constante atualização das práticas de moderação, *lato sensu*, para se adaptarem à legislação americana, às decisões da Suprema Corte Americana, à responsabilidade corporativa e principalmente à necessidade econômica de corresponder às expectativas dos usuários. As empresas gestoras dos fluxos comunicacionais têm sede nos EUA e, a rigor, seguem as orientações do seu país de origem.

No estudo a autora torna claro que existem um conjunto de forças que influenciam no comportamento das empresas de tecnologia para a moderação do conteúdo, *lato sensu*, e investem no aperfeiçoamento dos algoritmos e em treinamento de pessoal para aplicação de um conjunto estabelecido de regras e procedimentos *ex-ante* e *ex-post* para adjudicação de conteúdo.

A moderação, lato sensu, é uma atividade complexa que teve início em razão de reivindicações nos EUA sobre violações de direitos autorais no ambiente virtual⁹, e que expandiu para outras temáticas como questões relacionadas a material de propaganda terrorista, pornografia infantil, violência explícita, fake news e discurso de ódio, por exemplo. Os vácuos legais abriram caminho a restrições à liberdade de expressão cada vez maiores dentro dos domínios privados online, no entanto, ainda que sejam propósitos bem-intencionados, a falta de regulação pode desencadear excessos nas políticas internas de remoção, bloqueio e restrição das manifestações prejudicando a diversidade de opiniões.

No Brasil, o WhatsApp (NÓBREGA, 2022) é o meio de informação mais consultado pela população e essa circunstância de ascensão das mídias sociais como vetor de comunicação em massa, não parece ser compatível com a imunidade de responsabilidade das empresas que fazem a gestão do trânsito de informações online. Apesar de serem postagens de terceiros, feitas por seus usuários, a permanência de conteúdo degradante, violento ou de natureza desinformativa traz riscos às pessoas, à sociedade e gera repercussões sociopolíticas e econômicas no mundo real.

Nesse sentido, o presidente-executivo do WhatsApp, Will Cathcart, declarou em entrevista (UOL, 2021) que ““Nós trabalhamos de forma proativa, sem qualquer regulamentação, para combater fake news” e acrescentou “Trabalhamos muito nos últimos anos para fazer coisas como detectar disparo de mensagens em massa e limitar o encaminhamento”, no entanto, a forma como procedem é obscura, limitando discursos de modo autoritário.

Na página do WhatsApp, na secção dos termos e serviços (WHATSAPP, 2021) há expressamente a indicação de que os serviços podem ser modificados, suspensos ou encerrados pelas razões abaixo apontadas: 1) se houver violação das disposições dos termos ou intenções de prejudicar; 2) colocar em risco ou expor juridicamente o WhatsApp, os usuários ou terceiros; ou 3) caso a conta não fique ativa após o registro ou permanecer inativa por um período prolongado podendo ser desativada ou apagada.

Apesar de não trazer de modo claro quais as situações podem ser alvo de moderação de conteúdo, lato sensu, no tópico referente a uso lícito e aceitável (WHATSAPP, 2021) explicita que os serviços devem ser acessados e utilizados somente para fins lícitos, autorizados e aceitáveis e, nesse sentido, elenca as seguintes hipóteses de uso proibido: a) de forma a violar, apropriar-se indevidamente ou infringir direitos do WhatsApp, dos usuários ou de terceiros,

⁹ Vide mais em: Seção 230 da Lei de Decência nas Comunicações dos EUA (ADC).

inclusive direitos de privacidade, de publicidade, de propriedade intelectual ou outros direitos de propriedade; (b) de forma ilícita, obscena, difamatória, ameaçadora, intimidadora, assediante, odiosa, ofensiva em termos raciais ou étnicos, ou que instigue ou encoraje condutas que sejam ilícitas ou inadequadas, como a incitação a crimes violentos, a exploração de crianças ou outras pessoas, a ação de colocá-las em perigo ou a coordenação de danos reais; (c) que envolvam declarações falsas, incorretas ou enganosas; (d) para se passar por outra pessoa; (e) para enviar comunicações ilícitas ou não permitidas, como mensagens em massa, mensagens automáticas, ligações automáticas e afins; ou (f) de forma a envolver o uso não pessoal dos serviços, a menos que esteja autorizado pelo WhatsApp.

A internet trouxe potência às pessoas e através dos mesmos meios tecnológicos os usuários têm reivindicado mudanças na concepção libertária da liberdade de expressão online e exercido influência para modificar o modelo de negócios que acaba por reger suas vidas. Como resultado, motivam ascensão de uma nova política regulatória de discursos. O Twitter, por exemplo, perdeu poder de atração de novos usuários em razão do ambiente tóxico e teve de traçar linhas gerais moderação de conteúdo. (NITRINI, 2021, p. 91).

Na página do Twitter (2022) há um comunicado sobre as regras para promover “o diálogo público global” e a justificativa de que essas regras poderão ser “alteradas de tempo em tempos para cumprir nossa meta de promover um diálogo público saudável”. A política de moderação, lato sensu, do Twitter inclui assuntos relacionados a manifestações violentas, terrorismo/extremismo violento, exploração sexual de menores, abuso ou assédio, conduta de propagação de ódio ou que versem sobre ataques violentos, suicídio e automutilação, nudez não consensual, produtos ou serviços ilegais, postagens que violem a integridade cívica, dentre outros temas, em que cada tópico há regras pormenorizadas sobre as situações proibidas.

As ações de fiscalização do Twitter ([20--]) implicam em medidas restritivas das mensagens como: rotular um tweet que pode conter informações contestadas ou enganosas, limitar a visibilidade do tweet, exigir por e-mail a remoção de um tweet especificado, identificando as políticas violadas; ocultar um tweet violador enquanto aguarda sua remoção, substituindo o conteúdo original por um aviso informando que o tweet não está mais disponível porque violou as regras. Há também ações relacionadas às contas como exigir edições de mídia ou do perfil para se adequar as regras tais como identificar e-mail ou número de telefone, colocar uma conta no modo somente leitura, verificação da propriedade da conta até a mais gravosa que é a suspensão permanente do perfil.

Existem ainda ações contra conteúdos que apesar de não violarem as regras são marcados com avisos como aqueles para identificar conteúdo potencialmente sensível que algumas pessoas podem não querer ver ou ainda para reter um tweet com base na idade em virtude de conteúdo adulto, considerando para tal qualquer mídia produzida e distribuída consensualmente que seja pornográfica ou destinada a causar excitação sexual. Além disso, pode também reter um tweet ou conta em um país se recebida solicitação válida e com escopo adequado de uma entidade autorizada, indicando claramente o conteúdo a ser retido.

Em janeiro de 2022, o Twitter lançou um comunicado (TWITTER, 2022b) a respeito de uma ferramenta em caráter experimental sobre de moderação de conteúdo, lato sensu, relacionado à desinformação, especialmente fake news tendo em vista o ano eleitoral no Brasil, no qual os usuários fariam denúncias, em tempo real, sobre informações potencialmente enganosas. Como se percebe, há um desenvolvimento contínuo de mecanismos para regular o discurso online e, tendo em vista que o ambiente digital é fluido e as modificações nos sistemas são operadas por ajustes algorítmicos, os usuários podem nem tomar conhecimento sobre novas políticas internas.

O Facebook e o Instagram têm divulgado relatórios mais minuciosos sobre a política de moderação, lato sensu, inclusive com a quantidade de exclusões e de mensagens e de contas e demonstrando por números que as medidas tecnológicas proativas têm sido mais eficientes em detectar conteúdo violador de suas políticas do que as denúncias. O Facebook (2022), por exemplo, publicou que detectou e removeu 97% do discurso de ódio antes que tenha sido denunciado.

Em sua página, Facebook divulga os Padrões da Comunidade (2022b) e declara que tem como objetivo criar um lugar em que as pessoas possam se expressar e tenham voz e que, apesar, do comprometimento com a expressão ser uma prioridade, reconhecem que a internet cria maiores oportunidades de abuso, por isso, limitam a expressão com base em um ou mais valores a seguir relacionados: autenticidade, segurança, privacidade e dignidade. Sobre a política de moderação estabelece como seções os temas relacionados seguintes conteúdos: 1) violência e comportamento criminoso, 2) segurança, 3) conteúdo questionável, 4) integridade e autenticidade e 5) respeito à propriedade intelectual. Cada uma dessas temáticas com subtópicos que explicitam as políticas de moderação, lato sensu, do Facebook e que limitam da liberdade de expressão.

No relatório do 4º trimestre de 2021, o Facebook (META, 2023a) por exemplo, elencou as seguintes políticas de moderação: nudez e atividade sexual adulta; intimidação e assédio;

ameaça infantil; nudez e abuso físico e exploração sexual; organizações perigosas; terrorismo e ódio organizado; contas falsas; discursos de ódio; bens regulamentados como drogas e armas de fogo; spam; suicídio e automutilação; violência e incitação; conteúdo violento e gráfico. De acordo com a temática é possível acessar e tomar conhecimento sobre números de conteúdo e contas afetadas.

Com relação à desinformação, o Facebook (META, 2023a) classifica o conteúdo de verificação dos fatos em falsos ou parcialmente falsos e, desde agosto de 2020, acrescentou dois novos rótulos “adulterados” e “fora de contexto” ao conteúdo que circula em sua comunidade. Quanto as contas falsas, no Relatório do 4º Trimestre de 2021 (META, 2022a), o Facebook informou que foram detectadas 1,7 bilhões de contas falsas. Em relação a COVID-19, o Facebook (2021) informou que removeu mais de 3000 contas que denominou de “superdisseminadores de desinformação de vacinas.”

No Instagram (META, 2023a) são consideradas violações os mesmos conteúdos que o Facebook considera. Em sua página na internet, expressamente declara nos Termos de Uso (INSTAGRAN, 2022a) a adesão ao modelo de negócios referido no item 2.1.1 em que o serviço é financiado pelo direcionamento da publicidade segundo dados pessoais coletados e informações sobre atividades e interesses do perfil do usuário.

Em tópico específico, denominado “Remoção de conteúdo e desativação ou encerramento de sua conta” o Instagram (2022a) declara que não controla os que terceiros fazem ou mencionam em seus espaços e se exime de responsabilidade a respeito das condutas ou ações, online ou offline, dos usuários ou pelo conteúdo postado por eles. Além disso, nas Diretrizes da Comunidade há textualmente a menção a exclusão de conteúdo, desativamento de contas caso o usuário não siga os Termos de Uso.

Além dessas redes sociais e dos critérios expostos, outras redes sociais também têm políticas de moderação, lato sensu, com critérios distintos. As restrições sobre discursos online se expandem, evoluindo com as métricas de triagem mediante uso de inteligência artificial e, desse modo, novos conteúdos são adicionados como objeto de controle. A título exemplificativo, são os controle sobre bullying digital e assédio com novos critérios de filtragem que o Facebook (2019) implementou no terceiro trimestre de 2021, inclusive com a ferramenta que notifica o usuário sobre o conteúdo ofensivo antes da publicação, dando a oportunidade para refletir e alterar a postagem de modo que o destinatário sequer receba o conteúdo prejudicial.

Cada rede social age segundo políticas internas diferentes, padrões e normas próprias o que aumenta ainda mais a gama de informações e perfis sujeitos a controle da liberdade de expressão em diferentes países e com idiomas distintos. Essas circunstâncias denotam a sofisticação das seleções algorítmicas que procedem filtragens considerando diferenças de linguagem e culturais. No contexto da pandemia, as moderações de conteúdo aprofundaram os mecanismos de seleção para conter desinformação relacionadas a Covid-19 e às vacinas, com o caso do Instagram (2020). Para o período das eleições o Facebook e Instagram anunciaram (META, 2022), no dia 15 de fevereiro de 2022, que terão canal de denúncias dedicado ao TSE, além de mecanismos para rotulagem de conteúdo e outras ações para evitar abusos.

É um processo contínuo de ajustes das políticas internas sobre a regulação de discursos mediante aditamentos dos termos de serviços que se sobrepõem indefinidamente e que variam no âmbito das redes sociais, com medidas cuja gradação varia desde o sombreamento das mensagens para diminuir a visualização, marcação com aviso sobre conteúdo, bloqueio de comentários e likes, suspensão temporária da mensagem ou perfil até a suspensão definitiva.

Não há um regramento uniforme, além disso o fluxo comunicacional global está concentrado em poucas empresas que gerenciam o trânsito de informações online e o risco é restringir de tal modo a liberdade de expressão, comprometendo a diversidade social e a pluralidade de fontes, até que se conclua um processo de dominação das pessoas reduzindo as escolhas em opções binárias como sim ou não.

A par de ser fluído e de fácil adaptação, o design digital apresenta contradições e desafios, problemas atuais como os limites da liberdade de expressão aplicados de maneira universal em diferentes culturas assim como a dissonância sobre a constitucionalidade de modelos de controle que não sejam judiciais permanecem pendentes de solução. O cenário belicoso das guerras de narrativas fundadas em fake news bem como a ameaça real do instrumentalismo, ou seja, o poder das empresas de tecnologia não apenas de automatizar o fluxo de comunicação, mas automatizar o próprio comportamento humano (ZUBOFF, 2020), faz crescer a importância da diretiva através de novos arranjos normativos para compatibilizar tecnologia com humanidade, lembrando os propósitos altruístas iniciais da internet.

2.3.2 Liberdade de expressão no ambiente online e instrumentalização

A aplicação dos direitos fundamentais no mundo digital tem sido um desafio, uma vez que as garantias não dependem unicamente de uma posição negativa ou positiva do Estado, mas da colaboração de outros atores que atuam na cena da comunicação moderna. Sociedade civil, poderes privados, gestores da comunicação global, governos nacionais e organismos internacionais, estão envolvidos no desafio de definir parâmetros para liberdade de expressão online. Todos com interesses próprios e muitas vezes divergentes entre si. (MACEDO JR, 2020).

Estar conectado não é mais uma opção, faz parte da vida moderna e as pessoas estão cada vez mais dependentes da internet e, nessa arquitetura da comunicação digital, os controladores dos fluxos comunicacionais que, apesar de serem empresas privadas, têm mais poder que muitas jurisdições. Nesse sentido, Timothy Ash (2017, p.40) constata: “As práticas operacionais internas, por vezes secretas, das superpotências privadas podem ser mais influentes do que as decisões dos legisladores e dos reguladores”. De igual modo, Loveluck (2018) se manifesta ao afirmar que a regulação da internet está sujeita a forte influência do design do hardware e software.

O Brasil, ao eleger a via judicial com ritos não propriamente céleres para enfrentar uma dinâmica veloz como a divulgação em ambiente online, coloca as empresas de tecnologia numa posição cômoda, em que não lhes são atribuídos os deveres de monitoramento do conteúdo em rede. Apesar disso, na prática, essas empresas já exercem supervisão nos fluxos, atuando de modo proativo em face de conteúdo ilícito ou mesmo contrário às expectativas dos usuários. Nesse sentido, para Ingo Sarlet e Ivar Hartman (2019) argumentam que, em razão da quantidade de manifestações diárias na esfera virtual, não se pode conceber que o Poder Judiciário continue a centralizar o controle de conteúdo online.

Os autores Gilmar Mendes e Victor Fernandes (2020) defendem que o modelo brasileiro não impede que as empresas de tecnologia procedam a exclusões de conteúdo tão logo tenham conhecimento de postagem que violem os termos de uso, a partir de notificação extrajudicial. Essa interpretação, ainda que não seja uma efetiva alteração do modelo de responsabilidade, parece reconhecer ou mesmo incentivar as exclusões pelas empresas gestoras dos fluxos comunicacionais, ainda que o Marco Civil da Internet (MCI), Lei 12.965/2014, no art. 19, não lhes imponha o dever de monitoramento prévio porque condicionou as exclusões de conteúdo e remoções de perfis a ordem judicial específica.

O direito, para regular a liberdade de expressão online, caminha para um modelo de convivência de normas técnicas com regulação estatal, filtragens algorítmicas e decisões

judiciais. É novo arranjo entre poder privado e poder público, tendo em vista o volume de dados online e a circunstância de que as empresas gestoras intermediadoras dos fluxos dispõem dos códigos operacionais para alterar os sistemas. Com a internet houve uma descentralização da produção da informação, mas a distribuição dos fluxos e a governança online a nível mundial se mantém concentrada nas “*Big Five*”.

Desse modo, houve uma reorganização dos poderes que fazem o controle do que circula online tendo em vista que os códigos operacionais estão nas mãos das Big Tech e disputam poder com os Estados. Nesse sentido, Timothy Ash (2017, p. 64) já expunha que as empresas gestoras dos fluxos comunicacionais tinham “mais capacidade para permitir ou limitar a liberdade de informação e expressão do que a maioria dos Estados.”

Também tratando sobre a perda da centralidade dos Estados Nacionais, analisa-se o rearranjo do poder político entre governos, superpotências tecnológicas, sociedade civil e organismos internacionais que disputam forças no ciberespaço, nos seguintes termos:

Esse rearranjo do poder político entre governos, instituições internacionais e fóruns multistakeholders revela que o modelo de governança da internet se afasta da predominância de uma autoridade central hierárquica qual ocorre dentro dos Estados Nacionais e se caracteriza, mais precisamente, pela formação de redes multilaterais em que atores independentes e operacionalmente autônomos se articulam reciprocamente. (MENDES; FERNANDES, 2020, p. 21).

Apesar da comunicação online ter tido um caráter revolucionário no sentido de que inverteu a lógica da comunicação passiva, viabilizando uma forma descentralizada de produção de conteúdo em que os usuários participam ativamente, a distribuição da informação permanece concentrada pelas Big Tech. Essas empresas cresceram sob um regime regulatório frouxo em que os usuários estão vulneráveis à deterioração de direitos fundamentais.

Para Paulo Brasil Menezes (2020) o ambiente online é fragmentado e como as informações não circulam livremente, ao contrário, são direcionadas ao perfil segundo preferências catalogadas do perfil, muitas vezes a pauta de discussão é determinada por robôs que impulsionam determinado conteúdo, comprometendo a pluralidade democrática e tornando a comunicação enviesada e artificial (MENEZES, 2020).

No livro *The Coast-Benefit Revolution*, Cass Sustein, ao fazer uma avaliação sobre custos e benefícios de direitos, revela os perigos que a restrição à liberdade de expressão em detrimento da democracia, sob o condão de defendê-la pode causar ainda mais danos à própria democracia.

Qualquer princípio de precaução, com respeito aos danos decorrentes da fala, poria a própria democracia em risco imediato e grave. A análise de custo-benefício pode

parecer muito melhor, mas sofre exatamente do mesmo vício, que é permitir que funcionários não confiáveis invoquem um padrão aparentemente neutro e abstratamente atraente em defesa de resultados que violam esse mesmo padrão. (SUSTEIN, 2018, p. 201). (tradução livre da autora)

Se os doutrinadores já expressavam a preocupação sobre a seleção de informação através de servidores públicos que a princípio perseguem fins públicos no desempenho de suas funções, o perigo se torna ainda maior quando as aspirações que motivam as filtragens são baseadas em aspectos econômicos e financeiros, como um projeto de manipulação das pessoas e consolidação do poder econômico privado na esfera política. Desse modo, a liberdade de expressão, como está organizada pelo modelo de negócios online, não é uma expressão democrática especialmente tendo em vista que Balkin(2018) classificou as empresas gestoras dos fluxos comunicacionais online como autoritárias.

A manipulação de comportamentos em escala ocorre pelo condicionamento da liberdade de expressão segundo o modelo de negócios online que, ao perseguir lucros, organiza a comunicação online segundo preferências do perfil, ao manipular informações através modulação e moderação do conteúdo que transita online reduz a pluralidade que chega ao usuário. Assim, primeiro ocorre a instrumentalização dos discursos e depois das pessoas. Isso significa que, usando o conhecimento armazenado e continuamente atualizado que detêm sobre os usuários, as empresas gestoras dos fluxos comunicacionais customizam a experiência online, prejudicando juízo de valor, inviabilizando juízo de fatos, induzindo de modo subliminar padrões de comportamento até corromper a autodeterminação dos usuários.

O projeto de poder não se esgota na automatização das informações direcionadas segundo análises algorítmicas, mas na automatização das vontades humanas, procedendo modificação comportamental mediante artifícios de persuasão psicológica. Esse poder é descrito por Shoshana Zuboff (2020) como instrumentalismo, segundo o qual os usuários não seriam mais tratados como seres de vontade própria, mas como meios para resultados lucrativos. Nesse aspecto, a pluralidade de fontes de informações e de opiniões se revela ainda mais relevante tendo em vista que maior diversidade reproduz melhor a realidade, fornecendo uma visão mais ampla completa sobre o contexto fático e aumentando o leque de possibilidades para as decisões políticas.

Como referido por Callejón (2018), o modelo de negócios não visa a reprodução da realidade ao ar livre, ao contrário, desconstrói a realidade compartilhada ao promover a conflituosidade e a polarização. Ao priorizar fake news e lançar dúvida além de fragilizar a confiança e as redes de cooperação entre os usuários, as pessoas acabam por se distanciar e se

dividir. Além disso, a verdade fática resta comprometida tendo em vista que as versões fantasiosas são privilegiadas e intencionalmente expostas de forma mais visível nas telas em virtude dos interesses econômicos. Há uma desconfiguração da realidade, nesse sentido, Eugênio Bucci (2019, p. 75) defende que os fatos têm sido substituídos por dados nas discussões políticas em que “aglomerados de cifras e tabelas traduziriam em dígitos frios a realidade quente”.

Para o autor “dados se impõem como fatos consumados – mas não são fatos, e sim representações de fatos, ou, mais ainda, representações sobre aglomerações de fatos” (BUCCI, 2019, p. 77), isso significa ainda nas palavras do autor que “dados são uma versão materializada dos fatos” (BUCCI, 2019, p. 78). Ele afirma que as discussões consideram mais números estatísticos, gráficos e pesquisas do que fatos, tornando-os remotos. Desse modo, o ambiente de discussão online oferece uma versão dos fatos que não corresponde à realidade pragmática, ou seja, há uma desconfiguração das experiências sentidas e vivenciadas ao representá-las em análises algorítmicas, desacoplando os sentimentos e a humanidade dos fatos.

Essa circunstância conduz a uma percepção equivocada do mundo real, em que “formas eletrônicas de representação acabam determinando os parâmetros para a formação de um juízo de valor dominante” (BUCCI, 2019, p. 78). O ambiente de discussão online não é neutro e nem espaço de liberdade, mas sim um projeto de dominação alienante que desconecta as pessoas da realidade fática e impõe de modo subconsciente a sacralização de dados estruturados segundo finalidades econômicas, manipulando escolhas e moldando comportamentos dos usuários.

Desse modo, a verdade fática é distorcida e manejada em expressões matemáticas, conduzindo a desconstruções, conspirando mesmo sobre aquilo que não paira dúvidas, desacreditando a política, a razão, a ciência, a história e a própria democracia. Harari (2020) enfatiza que a confiança foi o alicerce de toda organização social, que as redes de cooperação entre as pessoas eram ordens imaginadas, baseadas na crença de mitos compartilhados.

Nas sociedades modernas essas crenças compartilhadas como o ideal de justiça, a religião, os governos, as leis, linhas fronteiriças, sistema financeiro, signos linguísticos são abstrações baseadas na confiança mútua. Essas ordens imaginárias se materializam na vida cotidiana como autoridades religiosas, chefes de Estado, território de países, dinheiro, sistema de escrita e conduzem a objetivos comuns, escolhas políticas em prol da sociedade, processos democráticos e cooperação recíproca.

Segundo o autor (HARARI, 2020, p.41), a realidade imaginada repercute no mundo fático e “Uma vez que a cooperação humana em grande escala é baseada em mitos, a maneira como as pessoas cooperam pode ser alterada modificando-se os mitos – contando histórias diferentes”. Donde se pode inferir que a circulação de fake news, ao desconstruir a realidade compartilhada, disseminadas de modo sistemático por milícias digitais e em larga escala é danosa às redes de cooperação, tendo em vista que desarticula a organização social democrática e corrói a confiança nas crenças compartilhadas que conduziram ao progresso da humanidade.

A verdade factual segundo Hanna Arent (1967): “Conceptualmente, podemos chamar de verdade aquilo que não podemos mudar; metaforicamente, ela é o solo sobre o qual nos mantemos e o céu que se estende por cima de nós”, no entanto, a própria autora, reconhece que a verdade factual tem fragilidades e está sujeita a interpretações e opiniões, suscetíveis a juízos de valor. As fake news atuam justamente nesse intervalo, entre os fatos e as opiniões, moldando versões e desconstruindo fatos, fazendo ruir as certezas mais concretas.

Nessa esteira de sucessivas manipulações do conteúdo online, ora a liberdade de expressão é manejada sem limites sendo as fake news priorizadas, ora é cada vez mais contingenciada com expansão dos critérios que a impõem restrições . São ações contraditórias do ambiente online porque tanto pretendem atender as expectativas dos usuários quanto visam ao lucro e em muitos aspectos essas pretensões, de corresponder às expectativas dos usuários e ao lucro, são conflitantes. Nesse sentido, o manejo da liberdade de expressão, das informações e mesmo das fake news foi investigado diante de um projeto de instrumentalização da humanidade em que a submissão das vontades e dos comportamentos são operados por expressões matemáticas, ou seja, algorítmicos que não são visíveis.

2.4 Democracia e governança das redes

A governança da internet abrange um modelo participativo de outros atores além dos entes estatais, caracteriza-se por ser multistakeholders em que governos, sociedade civil, setor privado e comunidade técnica atuam na construção de padrões e princípios. No entanto, o multilateralismo não é um elemento vinculante para o conceito de governança tendo em vista que algumas decisões são tomadas por poucos ou apenas um ator (ROCILLO; DUARTE; ROMAN, 2021).

Desse modo, o modelo de governança da internet se organiza segundo uma nova dinâmica de forças e se distancia da “predominância de uma autoridade central hierárquica qual ocorre dentro dos Estados Nacionais e se caracteriza, mais precisamente, pela formação de redes multilaterais em que atores independentes e operacionalmente autônomos se articulam reciprocamente” (FERNANDES E MENDES, 2020, p. 21).

Nesse sentido, poder político e normas técnicas, como expressão de governos e empresas gigantes da área de tecnologia, disputam espaços na seara da regulação (ou desregulação) da internet. Os EUA exercem forte influência sobre o modelo regulatório especialmente porque as empresas de tecnologia têm sede no país e estão vinculadas aos processos políticos e jurisprudenciais americanos. Ainda que sob críticas e denúncias constantes de que os gestores da comunicação global modulam informações e procederem moderação de conteúdo de modo a influenciar decisões eleitorais, os EUA, por serem grande beneficiários dos dados pessoais mundialmente coletados durante a navegação, não procedem um controle dos fluxos, adotando uma postura mais próxima ao determinismo tecnológico.

2.4.1 Governança das redes, liberdade de expressão e regulação

A compreensão americana sobre a internet como um processo natural e espontâneo fecha os olhos às repercussões econômicas, políticas e sociais nos espaços públicos e decisões democráticas. Num estudo, Alessandro Morelli e Oreste Pollicino (2020) fazem uma comparação de julgados sobre direitos humanos relacionados à liberdade de expressão, nos EUA e na Europa, entre a Suprema Corte Americana e o Tribunal Europeu de Direitos Humanos (TEDH), e concluem que a compreensão que se faz das redes interfere no resultado do julgamento, ou seja, a forma como os governos e tribunais compreendem a internet faz divergir as decisões judiciais.

Segundo os autores, nos EUA, no julgamento *Reno vs ACLU41*, em 1997, adotou-se a compreensão de que a internet é o “novo mercado de ideias”, um espaço de livre circulação, aludindo a metáfora do *Justice Holmes Jr.*, feita, em 1919, no julgamento do caso *Abrams vs United States*. Como se vê, a Suprema Corte reitera a mesma postura manifestada 80 anos antes. Na Europa, há uma “desconfiança substancial na nova tecnologia” e a percepção de que a amplitude da divulgação online pode dar mais potência à liberdade de expressão na internet e comprometer outros direitos individuais de maneira mais gravosa.

Os autores enfatizam que a diferente percepção sobre a internet decorre da posição do observador que faz os julgamentos. A análise distante, de fora do mundo online, é feita como se o ciberespaço fosse um lugar separado do mundo, como ocorre nos EUA, embora seja impossível negar a influência dos bastidores políticos e econômicos. Distinta é a análise realizada pelo observador interno que procede à avaliação baseado na realidade empírica, fazendo a conexão entre o ciberespaço e o mundo real, considerando o sistema jurídico vigente, caso da Europa. Eles concluem que “A intuição que está por trás desse relato é que a escolha entre um ponto de vista interno e um externo nunca é uma decisão neutra, mas sim um movimento baseado em juízos de valor precisos.” (MORELLI; POLLICINO, 2020).

Os Estados Unidos, país sede da maioria das Big Tech, são beneficiários do modelo de negócios de extração e acumulação de dados tendo em vista que há um interesse estratégico das agências de segurança norte-americanas (CALLEJON, 2020) e isso explica porque a compreensão americana da internet se dá como observador externo, negando os bastidores políticos e econômicos sobre a realidade pragmática.

A Europa, de modo contrário, tem adotado um posicionamento de maior controle dos fluxos online sendo pioneira na edição de parâmetros legislativos e principiológicos sobre os usos e aplicações tecnológicas, a exemplo do Regulamento Europeu sobre Inteligência Artificial (UNIÃO EUROPEIA, 2021) e a Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seus ambientes (UNIÃO EUROPEIA, 2018).

Esses instrumentos internacionais visam ao resguardo de direitos humanos no cenário tecnológico, bem como princípios de transparência, equidade e segurança no tratamento dos dados e nesse sentido está mais próxima da construção social da tecnologia, cuja ideologia é sensível a mútua afetação entre inovação e contextos democráticos, onde os avanços tecnológicos são contidos ou estimulados à medida que servem a propósitos de consensos.

Um conceito amplo sobre governança da internet foi moldado pela ONU que reconhece a multiplicidade de atores, a interação entre normas jurídicas e técnicas como um ajuste entre as forças que concorrem no ciberespaço, a seguir transcrito:

Governança da internet é o desenvolvimento e a aplicação por parte dos governos, do setor privado e da sociedade civil, em seus respectivos papéis, de princípios, normas, regras, procedimentos decisórios e programas compartilhados que dão forma à evolução e uso da Internet. (ONU, 2005)

Diante do cenário tecnológico, da proliferação de normas técnicas, da opacidade dos códigos foi preciso estabelecer um novo parâmetro de regulação estatal, sendo ideal um modelo que permita as inovações sem subjugar o Direito. Nesse sentido, Juarez Freitas e Thomas Freitas

(2021, p. 55) ressaltam que: “Não se trata – sublinhe-se – de prestigiar o modelo regulatório exacerbadamente proibitivo ou refratário à inventividade dos ecossistemas de inovação. Intenta-se tão somente frear, em tempo útil, as inovações portadoras de impactos líquidos negativos”.

Não há como imaginar o ciberespaço sem regulação, ainda que não seja da forma tradicional por meio de leis impositivas. Sem romantismos Lawrence Lessig expõe que o código das redes é propriamente uma lei, que a arquitetura de *hardware* e *software* determina comportamentos e direciona escolhas políticas. Assim, a falta de regulação estatal sobre quais fins perseguir é preenchida por normas técnicas do *design*, norteadas por finalidades lucrativas.

Nossa escolha não é entre "regulamentação" e "sem regulamentação". O código regulamenta. Implementa valores, ou não. Ele habilita as liberdades ou as desabilita. Ele protege a privacidade ou promove o monitoramento. As pessoas escolhem como o código faz essas coisas. As pessoas escrevem o código. Portanto, a escolha não é se as pessoas decidirão como o ciberespaço regula. Pessoas - programadores - irão. A única escolha é se nós coletivamente teremos um papel na escolha deles - e, portanto, na determinação de como esses valores se regulam - ou se coletivamente permitiremos que os codificadores selecionem nossos valores para nós. (LESSIG, 2000)

Corroborando essa compreensão, Shoshana Zuboff (2020) revela que o capitalismo de vigilância prosperou nos vácuos legais, e só se tornou possível porque se lançou “[...] em espaços não mapeados da internet onde enfrentou poucos impedimentos jurídicos ou de concorrentes, com uma espécie invasora num ambiente sem predadores”. Apesar de ter perdido a centralidade estatal, mais que nunca o direito deve se ajustar e regular a nova dinâmica da comunicação online que interfere e molda o exercício da liberdade de expressão.

De outro modo, tornar a internet um campo de batalha livre, sem intervenção do direito é dar azo à violência argumentativa e à força do mais influente, porque num campo de intolerâncias – como atualmente se caracteriza o ambiente virtual – nem sempre os melhores argumentos sobressaem (SCHREIBER, 2020). A regulação é necessária tendo em vista que os espaços de debate públicos têm sido deteriorados pelo choque com o modelo de negócios da comunicação online (CALLEJON, 2018).

Há, como dito acima, uma relação verticalizada na qual as empresas de tecnologia têm poderes e possibilidades de controle sobre os usuários bem diferentes das leis governamentais ou determinações judiciais. Elas ditam regras por meio dos termos de serviços e julgam conflitos de direitos online ocorridos em seus domínios. Implementaram um modelo de comunicação cuja produção da informação é descentralizada e o acesso é, a princípio, democrático, mas a governança dos dados, ou seja, dos fluxos online, persegue lucros, onde

nem todos tem as mesmas oportunidades e nem tudo que circula é confiável. Nas redes, a publicação é livre porque, em regra, não é feita triagem prévia, no entanto, em momento posterior, as postagens podem ser removidas em razão de filtragem algorítmica ou denúncias de usuários.

Após publicar um artigo sobre a modulação e moderação de conteúdo pelas redes, Katie Klonick (2017) concluiu que o ambiente online intermediado por empresas de tecnologia que fazem a gestão da comunicação global não tinha uma governança democrática. Ela constatou que as plataformas selecionam ativamente conteúdo postado por seus usuários, mas as regras de como procedem são opacas e diferentes da política externa fornecida ao público. Desse modo, ela afirma que “o que torna uma cultura democrática, então, não é a governança democrática, mas a participação democrática”. (KLONICK, 2018, p. 1666).

Assim, pode-se dizer que a cultura da internet, como tem acesso, em regra, livre e aberto, tem uma participação democrática, mas a governança operada pelas empresas que fazem a intermediação dos fluxos comunicativos não é. Em verdade, até mesmo o acesso tem sido objeto de restrições, a exemplo do Instagram que nos Termos de Usos (INSTAGRAM, 2022) no tópico referente ao compromisso dos usuários, veda a participação na comunidade às pessoas condenadas por crime sexual.

Nesse sentido, a comunicação online, por estar nos domínios econômicos particulares das Big Tech que fazem a gestão dos fluxos comunicacionais a nível mundial, pode subjugar a liberdade de expressão e de informação tornando esses direitos suscetíveis à instrumentalização, por isso é importante resguardar direitos de modo sistêmico diante do desenvolvimento tecnológico. Apesar de parecer invisível, os códigos avançam sobre espaços democráticos atuando de modo paralelo, por essa razão é tão importante compreender o ciberespaço e os efeitos das fake news disseminadas online por milícias digitais em que a destruição da realidade compartilhada pode ser uma forma de dominação privada ainda que a governança online aparentemente não tenha um interesse político específico.

2.4.2 Diferença entre a governança da internet e a democracia considerando as características procedimentais de Bobbio

Ao tempo em que o ciberativismo pretende uma democracia direta online, por desacreditar nos representantes eleitos, nas instituições governamentais ou nas versões oficiais,

o público não se dá conta do poder tecnológico onipresente e invisível que interfere no fluxo comunicativo e que canaliza as insatisfações por mais democracia como ataques contra às instituições democráticas. Da forma como estrutura os fluxos, leva a crer que o desejo de mais democracia em outras esferas da sociedade civil equivale ao desejo de libertação da representação política.

A aspiração por mais democracia parece estar sendo direcionada contra a própria democracia, em que discursos de desconstrução ganham ressonância no ambiente online. Bobbio (2019, p. 69) argumenta que “não é nova essa exigência de que mais democracia se exprima como exigência de que a democracia representativa seja acompanhada ou substituída por democracia direta”.

Desse modo, não é propriamente a democracia representativa que está em crise, mas qualquer forma de organização social que tenha de conviver de maneira tão intensa com a força invisível que desarticula e desorienta as pessoas, sejam democracias ou regimes imperialistas. A China imperialista controla o fluxo de informações, ideias e imagens sob intensa supervisão política e, apesar de se declarar favorável a liberdade de expressão, impõe limites a circulação de manifestações online dentro de suas fronteiras (ASH, 2017). Nesse sentido, nem mesmo a democracia direta pode ser a solução para a situação atual, tendo em vista que não apenas o caos desinformativo desorienta, mas também o excesso de politização gera apatia política, conformismo das massas e repercute em involução democrática. (BOBBIO, 2019).

Para Bobbio (2019), o reflexo desse processo de ampliação dos espaços democráticos numa sociedade politicamente em expansão não seria da democracia representativa para a democracia direta, mas a passagem da democracia política para a democracia social, ou seja, a extensão do poder ascendente para outras organizações da sociedade civil.

Segundo o mesmo autor, nenhum sistema complexo como o Estado Moderno pode funcionar apenas com assembleia deliberativa dos cidadãos, sem intermediários, e com o referendun porque é inviável que todos os cidadãos participem ativamente de todas as decisões concernentes a eles. Assim, de acordo com Bobbio (2019), a representação política é indispensável à democracia moderna sendo impraticável a democracia direta em sociedade de massa e plural.

Isso ocorre porque democracia direta em sociedade de massa e plural não tem procedimentos definidos, as regras do jogo não foram estabelecidas, nem a força impulsiva que se concentra nos movimentos de ruptura tem se mostrado capaz de direcionar os rumos

civilizatórios, mesmo diante da facilidade de mobilização na internet. O que se vê no interior do ambiente de discussão online é uma degradação humana das relações de diálogo, de respeito recíproco e parâmetros de convivência harmônica. Callejón (2018) classificou como de involução democrática a transformação operada internamente pelo modelo de comunicação online.

Corroboram com essas conclusões sobre a desarticulação dos sistemas de cooperação mútua pelos algoritmos as seguintes práticas: compartilhamento de fake news nas guerras de narrativas, a tolerância ao discurso de ódio (SEDLER, 2016), a monetização das fake news (MOROZOV, 2020) e o estímulo a conflituosidade (CALLEJON, 2020).

Apesar da participação online ter solução técnica, em tese, para o grande número de votantes distribuídos em partes longínquas do território, a complexidade dos problemas enfrentados não pode ser simplificada pelos meios de transmissão. A tecnologia algorítmica com suas operações numéricas objetivas não capta todo o enredamento da linguagem e, ao fazer uma análise superficial dos dados atribuindo-lhe significados, tende a “[...] aplinar a imensa complexidade das relações humanas, simplificando narrativas complexas em regras algorítmicas concisas e explicações monocausais” (MOROZOV, 2020, p. 142-143). Portanto, é preciso cautela ao depositar expectativas de celeridade e certeza nas decisões algorítmicas apresentadas pelas Big Tech porque nem sempre a precisão matemática reflete o dinamismo social e nem a isonomia no tratamento dos dados espelha critérios de confiança, neutralidade e transparência nas análises.

Assim, de acordo com Bobbio (2019, p.35) democracia, em uma dimensão procedimental, se caracteriza com o “conjunto de regras (primárias ou fundamentais) que estabelecem quem está autorizado a tomar decisões coletivas e com quais procedimentos”. Dentro dos espaços em rede, nos termos do modelo de negócios do ambiente de discussão online, as práticas são obscuras e sigilosas, justificadas por normas técnicas em que compreensão não é acessível a todos e cujas políticas internas são voláteis para corresponder a fluidez dos interesses econômicos. A par das decisões afetarem os milhões de usuários, não são tomadas com base numa esfera pública que capta anseios da comunidade, mas guiados por finalidades lucrativas.

Para Bobbio (2019), a primeira regra procedimental da democracia é que o poder de tomar decisões seja atribuído a um número elevado de membros do grupo, a segunda refere-se a decisão por maioria e a terceira é que aqueles que são chamados a decidir estejam diante de alternativas reais e tenham condição de escolher por uma ou outra. Frise-se que, para o autor,

os consensos democráticos em sociedade de massa e plurais não equivalem a unanimidade tendo em vista que as decisões são por maioria. Além disso, para que as pessoas tenham condições reais para decidir, exigem-se garantias mínimas como liberdade de expressão, de associação, de reunião de modo que as situações que restringem essas garantias sejam bem definidas e previamente conhecidas.

Apesar da internet atingir e oportunizar muitas pessoas a participarem das decisões democráticas, coletando os resultados das votações em números mediante contagem dos “sim” ou “não”, as alternativas binárias não favorecem as decisões negociadas porque tornam difícil condensar milhares de manifestações e opiniões de uma sociedade de massa em alternativas reais para votação. Assim, a deliberação é falha, especialmente considerando que algumas pessoas têm uma visibilidade maior nas redes, contam com mais seguidores, a exemplo de artistas e celebridades, então a voz desses perfis tem alcance maior.

Além disso, nas democracias, o poder de voto e voz se equivalem, diferentemente do ambiente online. A estruturação da informação pelas empresas gestoras dos fluxos online conduz a pensamentos, comportamentos e conclusões dos usuários, manipulando de modo subliminar a vontade dos cidadãos e maculando as decisões democráticas. O ambiente online não é favorável para uma tomada de posição porque a modulação dos fluxos não oferece alternativas reais para formação de opiniões, especialmente tendo em conta a atividade das milícias digitais que constroem cenários artificiais para induzir as decisões democráticas.

Ainda que ambiente online detenha os meios tecnológicos, não operacionaliza decisões mediante a participação dos usuários, suas políticas internas são lastreadas por ganhos financeiros, mesmo iniciativas como as de moderação de conteúdo aparentemente altruístas estão mais atentas ao poder de atração de mais seguidores e as expectativas dos anunciantes do que em observância a direitos de terceiros.

Desse modo, os controles proativos dentro do modelo de negócios estão mais atentos a relação de perdas e ganhos, balanceando o interesse no número crescente de seguidores e na venda de espaços publicitários com a desinformação, ou seja, apesar do ambiente tóxico poluído por fake news não corresponder aos anseios dos usuários e anunciantes e causar perda do poder de atração de novos usuários e decréscimo dos anúncios, as fake news também geram lucros pelo engajamento que propiciam com compartilhamentos, comentários e curtidas.

Nesse sentido, Tim Wu (2012, p. 382) explica o efeito rede e como se torna mais poderosa à medida que mais pessoas aderem, “uma rede que todo mundo usa vale muito mais

que o valor somado de cem redes com o mesmo número de usuários da grande rede.” Assim, o ambiente tóxico estagna o número de usuários e foi essa circunstância que levou o Twitter a promover práticas de moderação de conteúdo em seus domínios. (NITRINI, 2021)

Ocorre que num contexto democrático ainda que a governança das redes seja autoritária e não tenha que corresponder as características procedimentais da democracia de Bobbio (2019), por ser uma empresa privada, a gestão dos fluxos comunicacionais afeta negativamente o ambiente democrático. O ponto de equilíbrio entre a desinformação e as filtragens algorítmicas proativas coincide quando um critério de moderação de conteúdo atende simultaneamente a uma avaliação de perdas e ganhos (custos e benefícios) e de perigo real e iminente.

Essa voluntária convergência é utópica porque os interesses são conflitantes. As redes lucram com a instabilidade, com a falsa sensação de iminente golpe ao tempo em que a democracia é minada porque as condições de diálogo colaborativo são diminuídas numa situação de tensão e crise, mesmo que aparente, por ter sido artificialmente criada e não corresponder à realidade.

Diante disso, a exclusão de fake news disseminadas por milícias digitais até agora não demonstrou inconstitucionalidade porque a princípio não revelaram interesse político imediato, no entanto, considerando as práticas dos modelos de negócios, fluidas e voláteis, esse critério deve ser regulado para que sejam impostas obrigações de imparcialidade e eficiência na moderação de conteúdo quanto às fake news disseminadas por milícias digitais.

As restrições à liberdade de expressão de terceiros, no caso de perpetradores de fake news em associação com outros perfis para disparo em massa, são constitucionais porque não existe um direito à desinformação que proteja fake news e lhes garanta a permanência online. Apesar disso, é necessário cautela quanto a precisão dos resultados das análises algorítmicas porque existem interesses opostos, dentro das próprias empresas que operam as filtragens, que conflitam com esse critério de controle de fluxos e, por essa razão, a regulação estatal não pode ser desprezada no intuito de aperfeiçoar e impor obrigações e responsabilidades para evitar distorções futuras.

A arquitetura do ciberespaço é uma estrutura fluida que rapidamente se molda as influências externas ao mesmo tempo em que consegue desviar dos entraves jurídicos, expandindo nos vácuos legais. Desse modo, o contínuo desenvolvimento de regras internas para adaptação do modelo de negócios têm diversificado as práticas online que ainda assim

permanecem obscuras. Toda experiência humana em rede, em escala global, é quantificada em dados, armazenados em bases de dados privados, muitas vezes secretos, não oferecendo transparência às ações das empresas e às análises algorítmicas.

Captar aspirações democráticas com base em dados matemáticos apropriados sem consentimento em que os debates do ambiente de discussão online não têm procedimentos definidos para participação popular, demonstra como o ciberativismo não equivale a democracia. Ainda que o meio online tenha o potencial de penetração social profundo, o modelo de negócios online não é propício a uma democracia, nem mesmo quanto a sua própria governança.

Os usuários não têm poder de escolha nem autodeterminação sobre seus dados, muitas vezes captados sem consentimentos, e as escolhas pessoais manifestadas online não foram mediante opções reais dos usuários porque até mesmo suas vontades estão condicionadas por artifícios eletrônicos subliminares operados por expressões matemáticas frias. A aprovação da EC nº 115, no dia 10 de fevereiro de 2022, que incluiu a proteção de dados pessoais, inclusive nos meios digitais, como direito fundamental inserido no art. 5º, LXXIX da CF/1988, com o escopo de restringir o extrativismo de dados no ambiente digital e resguardar direitos dos usuários.

Desse modo, se o ambiente de discussão online não se enquadra como esfera pública nos termos do conceito de Habermas, porque a governança na internet é autoritária, sendo, portanto, uma arquitetura artificial em domínios privados com finalidades lucrativas em que a cláusula geral de liberdade de expressão e as medidas reparadoras constitucionalmente previstas e que foram construídas na perspectiva um ambiente democrático não se mostram compatíveis com o modelo de negócios que domina o ciberespaço.

O ambiente online não tem por objetivo reconstituir a realidade social, não há uma correspondência simétrica entre real e virtual ou um compromisso com a verossimilhança entre os mundos, por isso se justificam regras distintas entre a democracia e a governança digital, tendo em vista que são espaços diferentes com finalidades diversas. Diante dessas conclusões é possível conceber maiores responsabilidades para as empresas tendo em vista as práticas operacionais online que modulam e moderam a liberdade de expressão e de informação no ambiente virtual.

Assim, apesar de não previstas constitucionalmente, as exclusões de fake news disseminadas por milícias digitais são medidas de controle necessárias para o ambiente online,

e a princípio não são vedadas, desde que operadas de modo imparcial, com eficiência e transparência e, desse modo, podem diminuir a circulação de conteúdo desinformativo. No entanto, o poder público, por estar submetido as regras constitucionais, tem outro enquadramento jurídico.

Os potenciais gravosos da liberdade de expressão e de informação são amplificados pelas dimensões de tempo e espaço da internet em que a velocidade de disseminação e o alcance de divulgação expande exponencialmente os danos por isso modelos de controle célere e sistêmicos são importantes, contudo, devem estar sujeitos a obrigações e responsabilidades estabelecidas em lei para evitar mau uso e arbitrariedades.

3 DESINFORMAÇÃO ONLINE

Imersa na estrutura comunicativa de mercantilização da liberdade de expressão em que as manifestações pessoais são apropriadas e os dados adquiriram valor econômico, em que as mensagens são monetizadas a partir da quantidade de curtidas e compartilhamentos, em que o conteúdo importa menos do que os lucros para circulação viral das informações e notícias, as fake news adquiriram um status lucrativo. No ciberespaço fatos, opinião, razão, emoção, informação e desinformação transitam sem distinção como se fosse irrelevante o que diferencia verdadeiro e falso, ciência e ficção, fatos e versões, tornando difícil discernir realidade empírica da desconstrução artificial e direcionada das percepções.

A tecnologia está em constante desenvolvimento e, se antes a esfera digital era influenciada pelo mundo da vida, houve uma inversão desse fluxo tendo em vista que é o mundo virtual que parece moldar as escolhas dos indivíduos, apontar as predileções, comprometendo a diversidade de opiniões e a pluralidade democrática através mecanismos de personalização da experiência de navegação.

Ao agrupar perfis e segmentá-los segundo padrões de comportamentos são fabricados cenários para diferentes públicos de modo que realidades paralelas coexistem, cada qual com sua versão particular sem uma visão do todo, reforçando concepções e radicalizando opiniões. Assim sendo e ciente de que o objetivo do modelo de negócios do ambiente de discussão online não é informar, o conteúdo circula de acordo com finalidades lucrativas e não segundo fins de entendimento ou interesses públicos da comunicação numa esfera pública.

As práticas de modulação e de moderação de conteúdo online tornam as empresas de tecnologia muito mais do que meras intermediadoras da comunicação online, mas gestores do trânsito de informações, direcionando e manipulando as manifestações dos usuários que circulam em seus domínios. Uma abordagem mais compatível com o modelo econômico online realça a importância da regulação da internet no intuito de obrigações de transparência e responsabilidades das empresas de tecnologia em virtude dos efeitos nocivos que suas atividades provocam, tanto para as vítimas quanto para interesses coletivos juridicamente relevantes.

A par desse cenário, se as fake news não são novidades, é indubitável que a complexidade online trouxe consequências sem precedentes para esse fenômeno. A amplitude de divulgação com a transposição de fronteiras torna o alcance desterritorializado impondo

soluções além dos horizontes nacionais, a velocidade de disseminação trouxe a emergência de novas formas de controle deixando evidente que a dinâmica da internet é muito mais célere que os trâmites judiciais.

As medidas operadas pelas empresas têm se mostrado mais efetivas do que as decisões dos tribunais. Some-se ainda a circunstância de que a memória virtual é longeva a uma perspectiva etérea e, assim, enquanto não excluídas as fake news, continuam a propagar efeitos tendo em vista que não são esquecidas nem amareladas pela ação tempo, permanecendo vívidas.

Além das dimensões de tempo e espaço expandidas, o aperfeiçoamento tecnológico conduz a um incessante processo de inovação, razão pela qual são exigidas constantes adaptações das pessoas, das instituições, dos governos, das leis e da atuação judicial. É nesse quadro, que nascem novas medidas de defesa e combate às fake news e de enfrentamento às suas múltiplas exteriorizações no sentido de dar mais visibilidade as informações verificadas, balanceadas e verdadeiras.

3.1 Fake news

A comunicação online horizontal, interativa e de reciprocidade quase simultânea forjou um novo homem, capaz de transpor a realidade física para o modo virtual. Conviver com a dualidade de mundos, físico e digital, trouxe também novos desafios, dentre eles distinguir fato e opinião, verdadeiro e falso, informação e entretenimento. Desde a versão 2.0 da internet, em que usuários são produtores e consumidores de informação, não há uma checagem prévia dos fatos nem seleção das publicações, houve uma explosão de fake news no ambiente digital, sendo um efeito colateral da ampla conectividade. Em virtude disso, verdade, mentira, informação, humor, fato e opinião são publicados sem se que diferenciem os conteúdos.

Desse modo, identificar uma fake news pode não ser uma tarefa simples tendo em vista serem postagens com aparência de veracidade, que muitas vezes se apresentam como matéria jornalística e trazem carga emocional que despertam reações como raiva, medo e indignação cujo propósito da confecção/produção é enganar e/ou aferir lucro pela circulação viral. Nesse sentido, Juliano Maranhão e Ricardo Campos (2020, p. 322) entendem fake news como “notícias fraudulentas ou notícias falsificadas, ou seja, o conteúdo falsificado como jornalismo, nas diferentes mídias, com potencial lesivo”.

No mesmo sentido, Diogo Rais entende que nas fake news há uma intencionalidade em prejudicar ou obter vantagem. O autor reconhece que o significado diverso do termo inviabiliza o diagnóstico, mas traz uma definição aproximada para o Direito em que fake news podem ser identificadas como “uma mensagem propositalmente mentirosa capaz de gerar dano efetivo ou potencial em busca de alguma vantagem” (RAIS, 2020, p. 251).

A caracterização do ânimo fraudulento do emissor de causar prejuízo ou aferir vantagem, é difícil comprovar porque segundo Bolzan de Moraes e Adriana Festugatto (2021, p. 91) a exigência de comprovação de má-fé do produtor da notícia “torna quase impossível a persecução de um possível responsável, já que a comprovação do nexo de conhecimento sobre a falsidade da informação divulgada pertence, de regra, ao mundo íntimo do agente”.

Nem mesmo o dever de verdade é pacífico. Segundo Sarlet (2020, p. 515), não existe um dever de verdade quanto aos fatos nem quanto as opiniões, e inclui nessa seara fatos históricos, apesar de reconhecer que a negativa do holocausto tem gerado responsabilização na Alemanha. Note-se que o holocausto é um fato histórico comprovado e não uma opinião ou ideia. Bobbio (2020, p. 33) percebe uma tendência contrária ao direito de expressar as próprias opiniões sem o cuidado de checar previamente as informações “para que o homem não seja enganado, excitado ou perturbado por uma propaganda maciça e perturbadora”. De acordo com Bobbio (2020), há uma precaução ao exercício da liberdade de expressão diante das inovações na comunicação social online tratada como “maciça e perturbadora”.

Nesse sentido, Tomazzo Venturinni (2019) critica a separação entre o falso e verdadeiro e considera problemática essa dicotomia porque a oposição binária nunca é direta, para o autor a distinção que vale a pena fazer não é entre informações fabricadas e não afetadas, mas entre histórias que são apoiadas por um grande e honesto trabalho de fundamentação da verdade e histórias que não são. Ainda segundo ele, a desinformação não tem a intenção de enganar e de persuadir com a adesão cognitiva, mas de atingir a atenção.

Para Callejón (2022a) a construção social da realidade é uma obra coletiva que atribui significado aos fatos e condiciona a percepção do que se considera verdade. Segundo ele, na democracia pluralista há a necessidade de considerar as verdades dos outros a partir da transação, do consenso e do compromisso para conhecer a verdade em toda a sua dimensão. Ocorre que, segundo o autor, a estrutura comunicativa online está destruindo essa percepção conjunta da realidade tendo em vista que ao potencializar fake news os algoritmos as convertem em tendências majoritárias e assim alteram as condições estruturais dos processos democráticos.

Ao examinar a possibilidade de exclusão de fake news disseminadas por milícias digitais, não está se impor o dever de verdade nas manifestações online, mas o estudo se orienta por um mecanismo tecnológico que proceda moderação de conteúdo sistêmico e em tempo real que permita maior circulação da informação verificada e balanceada. Não se está a buscar formas de restrições da liberdade de expressão, mas recursos para que as informações que circulam online sejam mais confiáveis, então encontrar um critério de contingenciamento das fake news pode abrir espaço nos feeds para notícias verificadas, de fontes éticas e profissionais e ainda minimizar ou impedir a produção dos efeitos danosos das fake news disseminadas por milícias digitais que por terem um espectro ampliado, ao atingirem muitos perfis rapidamente, são mais abrangentes.

Desse modo, não se está a buscar uma proibição para o usuário se manifestar, inclusive, nos casos em que há a dificuldade de traçar uma linha divisória entre opinião e fatos falsos e fraudulentos, é preferível que a publicação permaneça online. De igual forma, o compartilhamento desavisado de fake news sem os cuidados médios de checagem ou os erros de avaliação sobre os fatos não estão sob discussão no trabalho, o recorte são as fake news disseminadas por milícias digitais.

Manifestações culturais e crenças religiosas também não se enquadram como fake news ainda que seja impossível comprovar a verdade factual. Nesse sentido, é garantida a liberdade religiosa de quem professa a fé e, por exemplo, acredita e divulga que o mundo foi criado em 07 dias. No mesmo sentido, também não se caracterizam como fake news lendas urbanas ou folclóricas.

Sob o aspecto do conteúdo, as fake news não são necessariamente fatos falsos, podem ser informações verdadeiras incompletas, descontextualizadas, relativas a fatos pretéritos, mas apresentadas como algo novo/atual, ou ainda enviesadas, por estarem relacionadas a pessoas que não tem nenhuma ligação com o acontecimento, por exemplo. Podem ainda, sendo informações verdadeiras, terem sido orquestradas para trazer uma falsa impressão da realidade como retrato de um quadro geral que só apresenta bons índices e ignora propositalmente outros números que complementam a representação mais fidedigna. Portanto, o que determina fake news é uma avaliação conjunta do conteúdo e do contexto.

Como se percebe, as fake news são um fenômeno complexo, de natureza fraudulenta, mecanismo para fisgar a atenção dos usuários online e capaz de destruir a realidade compartilhada de modo que identificá-las exige um juízo de interpretação, uma consciência crítica. A educação midiática pode contribuir para diminuir o compartilhamento, no entanto, a

impulsividade dos cliques numa guerra de narrativas ou segundo emoções despertadas pelas fake news torna a disseminação muito mais reativa do que racional.

Nesse sentido, Chiara Spadaccini de Teffé e Carlos Affonso Souza (2020, p. 290) comenta que “O Estado poderia, por exemplo, desenvolver campanhas de esclarecimento e combate às fake news, qualificar professores e determinados profissionais para tratarem do assunto em sala de aula e em outros locais estratégicos e investir em projetos voltados à Internet and media literacy.”

Apesar disso, o Presidente da República, em janeiro de 2023, vetou o art. 7º, § 11, da Lei 14.533(BRASIL, 2023c), que instituiu a Política Nacional de Educação Digital quanto a inclusão da educação digital com foco no letramento digital e no ensino de computação, programação, robótica e outras competências digitais no currículo do ensino fundamental e médio, sob a justificativa de que a inclusão de novos componentes curriculares depende de aprovação do Conselho Nacional de Educação e de homologação pelo Ministro da Educação.

Dessa forma, quanto ao enfrentamento da desinformação, a questão não é a incapacidade humana de autocontrole que potencializa as fake news, mas a própria dinâmica online que estimula comportamentos de forma inconsciente. Os usuários competem com estruturas algorítmicas que usam inteligência artificial para direcionar conteúdo segundo o perfil de cada um no intuito de despertar-lhes emoções que embarçam o raciocínio. Desse modo, há uma desproporcionalidade entre a capacidade humana de processamento e a tecnologia que direciona e estrutura a informação de modo a induzir as ações dos usuários.

As fake news podem ser entendidas como a versão eletrônica dos boatos, mas os efeitos se diferenciam porque, no meio virtual, o alcance é bem maior tendo em vista que a visualização não se restringe aos compartilhamentos entre os usuários. Os algoritmos amplificam a audiência de acordo com a quantidade de likes, comentários e compartilhamentos, além disso, no ciberespaço os efeitos são mais expandidos pela escala de contatos online que tende a ser maior por ser mais superficial, além da velocidade da comunicação digital e das análises algorítmicas com o uso de IA que direcionam de modo automatizado conteúdo de acordo com preferências do perfil.

Em maioria, as fake news circulam nas redes sociais e serviços de mensageria privada¹⁰ no contexto de conversas familiares ou entre amigos misturadas a informação e entretenimento

¹⁰ Tem-se o conceito normativo: “Mensageria privada é provedor de aplicação que prestam serviços de mensagens instantâneas por meio de comunicação interpessoal, acessíveis a partir de terminais móveis com alta

e é difícil generalizar um conceito preciso sobre fake news, porque são variadas as formas de manifestação, múltiplas as finalidades e com uma infinidade de assuntos abordados em seus conteúdos, circunstâncias que dificultam a formação de um conceito universal. Nesse sentido, Cass Sunstein ao tratar sobre boato diz que:

Não existe uma definição clássica para boato, e não tentarei oferecer uma aqui. Para dar início à discussão, devemos reconhecer a imperfeição de qualquer definição, abandonar as discussões semânticas e usar o termo para fazer referência, *grosso modo*, a alegação de fatos - sobre pessoas, grupos, acontecimentos e instituições – que ainda não foram comprovados, embora passem de uma pessoa para outra e, portanto, tenham credibilidade não porque se conheçam evidências diretas para corroborá-los, mas porque parece que outras pessoas acreditam neles. (SUNSTEIN, 2010, p.07)

Passados mais de 10 anos, a dificuldade de definir boato acompanha também a sina das fake news, que podem ser apresentadas como a versão tecnológica dos boatos. Por não ter um conceito consolidado, fake news é uma expressão usada corriqueiramente, com contornos indefinidos e com diferentes manifestações como *deepfakes*¹¹, *shallow fakes*¹² e *junk fakes*¹³.

As fake news atingem não somente direitos individuais que colidem com a liberdade de expressão e que disputam qual prevalecerá num caso concreto. Fake news também afetam “interesses coletivos juridicamente relevantes”, termos em que Wolfgang Hoffmann-Riem (2021) usa para se referir a democracia, saúde pública, família dentre outras instituições. Como esses são conceitos abertos, de conteúdo não definido e expressões amplas fazem rememorar a insegurança jurídica descrita como “fuga do juiz” em que a interpretação de modo elástico ocasiona a instabilidade jurídica e social. A medida da liberdade de expressão diante da democracia é uma linha tênue e as fake news testam esses limites.

Desse modo, a compreensão da democracia, por exemplo, fora dos propósitos nacionais causa distorções à proteção do ideal democrático, podendo ensejar tanto limitações do discurso de maneira arbitrária quanto a defesa da liberdade de expressão irrestrita, como direito individual absoluto que legitima fake news. Nesse sentido, Gundim e Pereira pragmatizam o paradoxo da democracia.

capacidade de processamento ou de outros equipamentos digitais conectados à rede, destinados, principalmente, à comunicação privada entre seus usuários, inclusive os criptografados.” (Art. 4º, XI, do PL 2630/2020)

¹¹ *Deep fake* significa “um modo profundo e refinado de propagação da desinformação por meio de vídeos em sua grande maioria de personalidades, cuja imagem e áudios são adulterados com o intuito de manipular a população” (RAIS, 2020, p 252)

¹² Paulo Brasil Menezes (2020. P. 118) explica que “*shallow fakes* utilizam aparatos de edição de vídeo propriamente dito para inserirem arquivos descontextualizados, mas sem haver qualquer alteração de conteúdo.”

¹³ Tomazzo Venturini (2019) *junk news* é perigoso não porque é falso, mas porque satura o debate público, deixando pouco espaço para outras discussões, reduzindo a riqueza do debate público e evitando que notícias mais importantes seja ouvidas. São notícias que perturbam, que desviam a atenção.

Eis o paradoxo da democracia: se, de um lado, a liberdade de pensamento e, mais especificamente, a liberdade de expressão constitui ponto fulcral para sua preservação e para o seu desenvolvimento, o exercício de tal liberdade por agentes mal-intencionados põe em risco a própria democracia. (GUNDIM; PEREIRA, 2020, p. 260).

Portanto, abordar fake news implica tratar também sobre seus efeitos para a democracia, liberdade de expressão, de informação e pluralismo. Como não é possível trazer uma definição exata das fake news torna-se importante investigar o fenômeno ainda que seus contornos sejam fluidos e estejam em contínua evolução, e mesmo sendo inexitoso esgotar as características, o estudo pode abordar algumas finalidades, temáticas, razões pelas quais prosperam, efeitos e formas de controle para que se tenha uma visão geral.

Assim, sobre as finalidades das fake news, Paulo Brasil Menezes (2020, p. 130-136) traz num rol não exaustivo sobre as finalidades, dentre as quais a: 1) finalidade separatista que promove a segregação ou alienação o povo a medida que lhe retira a capacidade de se articular para mudar a realidade; 2) para estimular “o populismo cibernético em que a racionalidade política tem cedido lugar para a emoção efêmera humana”; 3) ou para fabricar consensos políticos a partir da intromissão de notícias fraudulentas de modo contínuo no intuito de “desconfigurar a percepção opinativa dos cidadãos”;

O autor prossegue explicando outras finalidades como as de: 4) reforçar o subjetivismo das pessoas pela repetição de notícias direcionadas inclusive fake news fragmentando a racionalidade pública e acirrando a polarização política; 5) solidificar as redes sociais, através da finalidade tecnológica, no intuito de robustecer o poder de influência da rede na esfera pública por meio artificial através de algoritmos que direcionam informações a perfis humanos ou manipulam a agenda política por perfis-robôs a fim de determinar a pauta de discussão; 6) vulgarizar o conhecimento técnico em que a tradução simplista de problemas complexos despreza estudo científico, bases teóricas e habilidades intelectuais mínimas necessárias para a compreensão das situações em detrimento de soluções superficiais.

Ainda restam três finalidades listadas por Paulo Brasil Menezes (2020) a seguir explicitadas: 7) segundo a finalidade estratégica abordada pelo autor, as fake news tem sido veiculadas para dar vazão a violência argumentativa, no intuito de estimular afrontas e liberar animosidades fazendo uso de linguagem extremista e acusatória das instituições democráticas, para alimentar emoções negativas e conteúdos inflamados e inviabilizar soluções democráticas; 8) a finalidade lucrativa corresponde a forma de remuneração das fake news que se dá pela quantidade de compartilhamentos e cliques. Assim, a incessante veiculação acarreta mais lucros e ampliação da rede de relacionamentos de quem propaga; e, por último, 9) a finalidade

dissidente em que a liberdade de expressão é feita para o transpasse da utilidade pública, ou seja, “da crítica positiva para a prejudicialidade da crítica negativa” (MENEZES, 2020, p. 160), favorecendo a desintegração do compartilhamento multidisciplinar das funções públicas necessárias para o espaço público.

As finalidades acima se tocam em muitos aspectos, sendo repercussões umas das outras, num processo que se autoalimenta com variadas técnicas de desinformação e que minam a confiança na democracia como regime capaz de promover o progresso social e as redes de cooperação entre indivíduos conectados. A globalização, a desterritorialização das dinâmicas modernas operadas por meio eletrônico, o declínio na confiança assim como a perda de credibilidade política dos representantes eleitos são situações que alimentam a crises democráticas do mundo interligado.

A temática das fake news é variada, não há um assunto-alvo, pode se referir a inventos e conquistas científicas, a conduta das autoridades públicas, a boatos sobre anônimos e famosos, falsas acusações, estar relacionada a situações emergenciais como crises sanitárias, a direitos de personalidade de pessoas comuns dentre muitas outras, mas foi na arena política que seus efeitos se tornaram mais poderosos.

O escândalo da Cambridge Analyticals nas eleições americanas de 2016, com interferências internacionais por meio de milícias digitais (BBC, 2022), alarmaram todo o mundo sobre o poder das mídias e a respectiva influência nas escolhas públicas, inclusive, políticas. Diante disso, começaram discussões em todo o globo sobre as práticas abusivas da comunicação eletrônica online. Vários enfoques foram trazidos à tona, dentre eles os limites da liberdade de expressão dos usuários, a prevalência das fake news em detrimento de informações verificadas e de fontes jornalísticas, a introjeção sublimar de significados pelas empresas de tecnologia, através da modulação e moderação de conteúdo.

Apesar de estarmos imersos na Sociedade do Conhecimento com grande quantidade de informação disponível por que as fake news prosperam e se tornam um fator de desagregação social e uma ameaça à democracia? Algumas circunstâncias podem trazer luz para a resposta da pergunta acima, a facilidade de acesso à internet, a postagem sem critério prévio de veracidade, a rede de contatos expandidos online ser maior do que a agenda de relacionamentos pessoais reverberando assim em mais pessoas, a produção e a viralização artificial por robôs, todos são fatores que impulsionam o alcance virtual, mas a lógica de mercado que remunera fake news e amplia a visibilidade é determinante.

Nesse sentido, Dan Wielsch (2020, p. 98) explica como ocorre a conectividade online: “Essas novas formas (por exemplo, ‘compartilhar’, ‘tornar-se amigo’ e ‘curtir’) facilitam a rápida formação de segmentos de audiência para atos de comunicação e referências correspondentes de outros membros.” Desse modo, as reações entre os usuários criam ligações em rede, sinalizando que a atenção foi catalisada com êxito, e passam a indicar pontos de referência para direcionamento de conteúdo e publicidade. Ainda segundo o autor, essa comunicação em rede, conectada por pontos, substituiu os vínculos sociais por ligações superficiais e implicou em aumento da quantidade de contatos comunicativos. Desse modo, os caminhos que uma mensagem online percorre são bem mais expandidos.

As fake news têm uma audiência expandida não apenas pelos padrões de conectividade, mas também porque o controle social na comunicação online é menos eficiente tendo em vista a forma de interação entre as pessoas. Nesse sentido, Barry Chudakov (2017) enfatiza que as reações corporais entre pessoas que conversam presencialmente podem conduzir o diálogo, alterando abordagens e procedendo autocontenções, de modo que um semblante em movimento afeta o desempenho da fala. Para ele, o diálogo presencial costuma ser conduzido por interações entre os atores, como um franzir de sobrancelha, um olhar, um movimento de cabeça, que sinalizam a conversa. No ambiente online, o máximo que há é um rosto congelado, um código em linha ou endereço e isso encoraja comportamentos desinibidos, sem freios, como guerra de narrativas fomentadas por fake news.

Desse modo, como no meio online a interação não é percebida da mesma forma que nos encontros presenciais e os vínculos sociais foram substituídos por contatos superficiais, a emocionalidade é exacerbada para fins de feedback, como uma curtida ou comentário e, assim, as narrativas apelam para sentimentos como táticas para fisgar atenção. E, como as fake news têm forte carga emocional para prender a atenção do público e gerar engajamento, essa pode ser uma das causas pelas quais são priorizadas na distribuição dos fluxos online.

Para Cass Sunstein (2010) os boatos prosperam por razões variadas, ele argumenta que a maior parte do nosso conhecimento é indireta, derivada dos outros e, por não termos certeza de tudo, há uma tendência a acreditar que as mensagens sejam verdadeiras, ainda que parcialmente, submetendo-se à sabedoria das multidões. Além disso, ele sustenta que as convicções pessoais prévias reforçam a adesão aos boatos porque as pessoas estão predispostas a acreditar em algo compatível com seus interesses ou com o quê acreditam ser verdade, assim, também acreditam em boatos em razão dos medos e esperanças particulares.

Nesse condão de explicar a difusão acelerada dos boatos, Cass Sunstein (2010) argumenta que dois fatores, as cascatas sociais e a polarização de grupos, potencializam a circulação. Ele explica que a falta de certeza leva as pessoas a confiarem em alguém ou alguma instituição que cada um atribuiu como líder e que arrasta seguidores. Desse modo, como primeiro fator que ele denominou como cascatas sociais, ele atribui ao comportamento de líderes de acreditar e divulgar boatos capazes de produzir especulação provocando o efeito em cadeia.

Como segundo fator, ele enumera a polarização de grupos e se refere a circunstância de que uma crença hesitante individual se torna certeza a partir dos contatos com grupos de pessoas que pensam de forma semelhante. Nesses casos, as trocas entre pessoas que pensam de forma semelhante tornam os argumentos mais radicais e menos permeáveis aos debates. O autor explica que isso ocorre até mesmo como forma para preservar a auto reputação, na medida em que se nega a oportunidade de mudar de opinião para manter a imagem de que está correto perante outros membros do grupo. Diz ainda que as emoções despertadas ou inflamadas obstruem o caminho da verdade tendo em vista que as pessoas não processam as informações com neutralidade. Essas ilações podem igualmente se aplicar às fake news.

As fake news ganham repercussão no intervalo entre os acontecimentos e as opiniões, no processo de internalização dos fatos e interpretação das situações. A linguagem escorregadia alimenta especulações e abre caminho a subjetivismos, convertendo o terreno dos fatos em espaços de exercício imaginativo. Assim, a instabilidade de interpretações fundadas em falsidades, emoções ou em informações incompletas ou descontextualizadas torna a leitura das situações distorcida, induzindo uma compreensão fantasiosa dos fatos, inclusive de forma subliminar, prejudicando a avaliação da realidade, maculando as tomadas de posição individuais e manipulando a opinião pública.

De acordo com Vania Baldi (2018), os mecanismos tecnológicos de apropriação de dados pessoais formam perfis psicológicos não apenas de pessoas, mas de grupos populacionais inteiros e, assim, as fake news, dirigidas a um público previamente analisado, provocam um efeito arrebatador porque reforçam as concepções dos usuários, radicalizando as opiniões. Disseminadas para grupos previamente catalogados e receptivos, as fake news encontram campo fértil para se multiplicarem entre a rede de contatos, movimentando o ambiente de discussão e distanciando pessoas que pensam de modo divergente, colocando-as em posições opostas.

Diante das revelações sobre o modelo de negócios online, as fake news tendem a prosperar e se multiplicar no ciberespaço porque, na perspectiva econômica, “as fake news são apenas as notícias mais lucrativas” (MOROZOV, 2020, p. 11). Eugênio Bucci (2019) afirma que há um direcionamento algorítmico para potencializar os efeitos das fake news no ambiente online. Para ele, a soma de fatores como velocidade de propagação, alcance da divulgação em escala mundial e a eficácia da audiência selecionada tornam as fake news ainda mais nocivas por isso a disseminação de fake news por milícias digitais têm o potencial ainda mais destrutivo.

Pois bem, as fake news minam as redes de cooperação entre as pessoas a medida que as divide e as distancia, provocando a corrosão da credibilidade das comunicações públicas com ataques sistemáticos a instituições como a democracia, a saúde pública dentre outras, além de implicarem dos danos e ofensas às vítimas. As fake news são complexas, repercutem na liberdade de expressão e de informação, com consequências negativas para esferas individuais e coletivas.

Ao passo que se revelam as práticas comerciais da Big Tech, a partir de denúncias até mesmo de ex-funcionários, amadurece-se a necessidade de contingenciamento das fake news. A princípio, houve a percepção das empresas de tecnologia como meramente intermediadoras dos fluxos comunicacionais até a constatação atual de que são empresas de mídia social, gestoras de informações e reguladoras da comunicação global com burocracias internas que criam e aplicam normas e valores, operando moderação e modulação de conteúdo online (BALKIN, 2020).

Assim, se avançou na percepção de que os controles praticados pelas próprias redes não são apenas direitos sobre seus domínios privados, mas devem ser internalizados como obrigações e responsabilidades dessas empresas tendo em vista os riscos que suas atividades representam para instituições, indivíduos e bens coletivos juridicamente relevantes como a democracia.

As formas de controle estatais atuais miram no comportamento das Big Tech (Balkin, 2018), que desvirtuam os propósitos da comunicação social, e em propostas que dificultem as práticas fluidas dessas empresas que rapidamente se reprogramam e escapam às regulações, mas não há uma medida infalível ou uma regulação definitiva. Apesar disso, é preciso desenvolver mecanismos de defesa contra o mar de desinformação, no qual estão inseridas as fake news. Essas estratégias não pretendem atingir a liberdade dos usuários que compartilham de maneira desavisada, sem os cuidados médios de checagem ou por erros de julgamento, no

entanto, existe a preocupação em fortalecer a verdade, dar mais visibilidade a informações verificadas e balanceadas e aperfeiçoar os filtros.

A abordagem de combate a desinformação está mais centrada sobre a atuação das empresas de tecnologia que dominam o ambiente online do que contra os usuários individualmente. As formas de controle das fake news necessitam de uma análise de conteúdo e contexto, diferenciando fato e ficção, verdadeiro e falso, razão e opinião, ciência e crença. Em muitos casos a linha que separa esses conceitos é tênue e a intenção informativa não está clara, é implícita, por isso modelos de moderação que associam controle humano e algorítmico, analisando conteúdo e atividade online podem ser mais eficientes.

No cenário político que atinge a cidadania, o pluralismo e a democracia, as fake news prejudicam a avaliação da realidade, direcionando a compreensão social e induzindo os usuários/cidadãos a erros de julgamento além de tumultuarem o trâmite regular das deliberações porque deturpam os fatos e prejudicam os juízos de valor, dificultando o entendimento já que as discussões não correspondem a situações verídicas, ficando as negociações tomadas por um movimento circular de argumentos baseados em informações imprecisas, descontextualizadas ou inverídicas seguidas do descortinamento das falsidades de modo que não progridem e não se ultimam as decisões. E nesse ciclo a cidadania e democracia se fragilizam e os extremistas ganham robustez e não se completa o ciclo de discussão, votação e deliberação tendo em vista que não se consegue avançar em virtude da polarização que prejudica a composição entre os interesses divergentes que se tornam cada vez mais polarizados e resistentes a concessões recíprocas.

Nesse sentido, Daniel Innerarity (2021, p. 172) é preciso: “A democracia é um sistema político que equilibra discussão e decisão, negociação e resolução, acordo e dissidência”. A democracia pressupõe deliberação, ou seja, ação de análise de propostas e votação para a formação ou não de uma medida legal compartilhada, em que foram oportunizadas e consideradas correntes ideológicas com interesses diversos, é, portanto, uma via de entendimento, um processo com começo, meio e fim.

O estudo das fake news envolve discussões que versam sobre a liberdade de expressão, de informação, dignidade humana, modelo de negócios online, modulação e moderação de conteúdo, controle de comportamento inconsciente, tecnologia, política, democracia, cidadania, pluralismo e esfera pública além de outras temáticas complexas com definições mutáveis segundo o tempo e o lugar. Em razão das mútuas interferências entre realidade e as dinâmicas online, com velocidade supersônica e em escala mundial, as medidas de contingenciamento das

fake news dependem de mecanismos que sejam rápidos e atravessem fronteiras territoriais. Portanto, ao longo do tópico, as fake news foram caracterizadas abordando conteúdo e contexto, tanto discorrendo finalidades, temáticas e efeitos sobre a democracia bem como a necessidade de contingenciamento dinâmico e sistêmico.

3.2 Fake news e a insuficiência/problemática do modelo de Moderação Online Judicial

No tópico 2.1.1, foi diferenciado o ambiente de discussão online da esfera pública, segundo o conceito de Habermas (2020), no entanto, as fake news disseminadas no ciberespaço deturpam os debates públicos. Diferentemente da esfera pública, cujo propósito é o entendimento, no ambiente de discussão online a conflituosidade é estimulada, a informação é direcionada e não necessariamente confiável, a liberdade de expressão individual não é igualmente distribuída em virtude da ação de disseminadores artificiais e a compreensão social é comprometida pelo uso de algoritmos que organizam as informações de modo tendencioso, introjetando significados.

Diante da intensa conectividade e participação online, os debates têm ocorrido, em maioria, em espaços virtuais, que são domínios digitais privados de acesso livre, a princípio. Ocorre que, esses ambientes são monopolizados por um modelo de negócios de poucos gestores que concentram o controle dos fluxos comunicacionais e a tecnologia para torná-los acessível. Com suas práticas, essas empresas tornaram a liberdade de expressão um produto comercial ao apropriarem das manifestações e comportamentos online dos usuários e os quantificarem em dados para análises algorítmicas posteriores, tornando o conteúdo online matéria-prima gerenciável e negociável segundo interesses privados.

De acordo com Francisco Callejón (2022a), além de condicionarem a liberdade de expressão dos usuários, determinando o público que atingirão, ampliando ou restringindo o alcance das manifestações individuais, essas empresas também estruturam as informações online para atrair atenção e gerar lucros com a venda publicitária, privilegiando conteúdo que gera mais engajamento como curtidas, comentários e compartilhamentos e assim, segundo essa lógica, radicalismos e fake news são potencializadas. No ambiente online, as repercussões das fake news são mais nocivas em virtude das dimensões expandidas dos ciberespaço, mas, apesar disso, a moderação de conteúdo padrão ainda permanece atrelada as competências

jurisdicionais e aos trâmites lentos. Há um desequilíbrio entre as dinâmicas online e os controles ao modo offline.

A comunicação online na versão 2.0 foi emancipadora para o indivíduo porque democratizou o acesso dando voz aos excluídos pela mídia tradicional, mas como disse Klonick (2017) a liberdade de expressão online é intermediada por governadores (=moderadores) do espaço digital razão pela qual não é tão livre quanto aparenta. Para Callejón (2022a) a gestão dos fluxos segundo interesses econômicos converte informação e opinião em mercadoria e, desse modo, a liberdade de expressão já não corresponde mais o sentido substancial de contribuir para o debate democrático e para pluralismo. Ainda segundo o autor, da forma como são manejadas as informações e opiniões dos usuários, a liberdade de expressão se torna um fator de agressão e radicalização e os efeitos da gestão dos fluxos online se materializam em interferências nos processos eleitorais e em involução democrática.

A própria governança dos dados é obscura, na medida em que os usuários desconhecem em quais grupos populacionais estão inseridos, o alcance de suas publicações, quais os assuntos que lhes são oferecidos com mais frequência, como e quando seus dados de perfis estão sendo negociados dentre outras informações relacionadas a dinâmica online. Para Giovanni Pitruzzella (2018) não está claro para as pessoas que as informações que lhes são dirigidas são dados seletivos, frações de um todo ou versões apontadas e fabricadas para um público que compartilha os mesmos interesses e preferências e que por isso podem induzir a falsa percepção de que um determinado ponto de vista é majoritário.

Assim a potência individual conferida pela liberdade de expressão em rede foi suplantada pelos interesses econômicos, as empresas de tecnologia desenvolveram burocracias sobre a gestão dos fluxos e se transformaram em empresas de mídia social, com enorme poder sobre sua comunidade e isso desarticulou os sistemas de controle democráticos pondo em xeque a proteção dos indivíduos diante de seus próprios pares e perante os gestores do modelo de comunicação online.

Nesse sentido, Balkin (2018) ressalta que as empresas de mídia social agem de maneira autoritária impondo suas regras sem transparência, modulando e moderando conteúdo, inclusive sobre fake news, julgando e aplicando sanções. Desse modo, coexistem ao menos duas esferas decisórias, a estatal e a privada com formas de solução diversas e, por dispor dos códigos operacionais, as empresas de tecnologia gestoras da comunicação online têm mecanismos mais eficazes para limitações da liberdade de expressão no meio digital, como a baixa visibilidade do conteúdo, o bloqueio temporário do perfil e exclusões automatizadas.

Ocorre que suas práticas são implementadas segundo critérios opacos. Timothy Ash se refere a dualidade de sistemas e as múltiplas composições que podem ser feitas entre ambos, nos seguintes termos:

Em suma, não existem dois países claramente delineados, com uma única fronteira nítida a separá-los: o Reino das Leis e a República das Normas. Mas também não existe uma continuidade suave entre a lei dura e a norma branda. Em vez disso, há uma série de descontinuidades à medida que se passa da lei penal mais firme (se disser isso, vai pra cadeia) para a norma mais privada (por favor, não diga isso, porque me perturba) (ASH, 2017, p. 106).

No âmbito virtual, as empresas que fazem a gestão dos fluxos comunicacionais estão, cada qual desenhando regras próprias, criando e as implementando para dirimir questões relativas à liberdade de expressão online. Gunther Teubner (2017, p. 195) argumenta que as constituições dos Estados-nação se mostram insuficientes e defende a criação de uma constituição digital que uniformize práticas no ciberespaço com a reconfiguração do constitucionalismo clássico baseado na relação entre indivíduo/Estado para uma relação coletividade/instituição.

Para Gunther Teubner (2017) o ponto mais crítico das dinâmicas online são os conflitos decorrentes das próprias dinâmicas digitais, não se trata apenas do poder econômico das empresas de tecnologia que é mais visível ou da posição de superioridade em relação aos usuários onde julgam conflitos em suas esferas, criam regras e as aplicam num sistema que escapa aos processos democráticos de elaboração e adjudicação, sendo produzidas e implementadas pelo mesmo centro de poder, sem divisão de função ou sistemas freios e contrapeso.

Para ele (TEUBNER, 2017), no ambiente digital, não basta uma proteção de direitos individuais dos usuários um frente ao outro ou perante as empresas de tecnologia gestoras dos fluxos comunicacionais porque existem movimentos em rede articulados em que os próprios usuários têm comportamentos direcionados em face de um alvo ou de um objetivo específico e que, apesar de se constituírem relações privadas internas, geram efeitos na sociedade e ameaçam instituições. Nesse sentido a atuação de milícias digitais é um movimento articulado em rede.

Ainda segundo o Teubner (2017), movimentos online despersonalizados e não formalmente registrados podem e devem ter uma regulamentação específica no ciberespaço porque tanto podem ferir direitos constitucionais individuais quanto ameaçar instituições, como os casos de associações de usuários que se organizam virtualmente em processos comunicativos autônomos e geram efeitos na sociedade, promovendo sanções sociais em escala, a exemplo dos

scores virtuais para ranqueamento de empresas condenando práticas comerciais sem oportunidade de defesa, manchando a imagem de suas atividades e repercutindo negativamente no desempenho econômico.

Nas situações de movimentos em rede, os indivíduos se relacionam como uma mentalidade coletiva em face de uma instituição ou de uma pessoa. A relação entre os usuários que se mobilizam, a princípio horizontal, toma força como estrutura social e como agente de transformação e, se mostra ao fim como uma relação de poder de muitos em detrimento de muitos ou de um só. Desse modo, agrupamentos de perfis por afinidades de interesses, projetos mútuos ou colaboração recíproca em movimentos online para fins ilícitos, como milícias digitais que disseminam fake news, inclusive com a articulação de perfis falsos e robôs, têm repercutido em relações sociais inéditas e de controle ainda incipiente.

Desse modo, a mera transposição de normas constitucionais para o âmbito digital colocando as empresas de tecnologia na posição de superioridade, equiparadas a figura estatal, não contempla as peculiaridades da estrutura digital, como as novas modalidades de organização social e as interferências do código nas regras comportamentais, por isso é crucial debater novas formas regulatórias sobre o fluxo online ciente desses novos arranjos sociais, da assimetria entre os usuários e as empresas de tecnologia que fazem a gestão dos fluxos comunicacionais online e da diversidade e dinamismo dos sistemas de moderação de conteúdos privados para contingenciamento das fake news.

Antes da comunicação social online o Estado era a grande ameaça à liberdade de expressão, no entanto, com a internet outras relações se formaram de natureza vertical como a superioridade das empresas de tecnologia em face dos usuários e a do poder social articulado em rede no qual a coletividade conectada mostra força diante de instituições, de modo que bens juridicamente relevantes como a família, a democracia e a religião podem ser alvos de organizações ou processos sociais autônomos.

As complexidades das relações entre os dois mundos, físico e digital, criam novas manifestações a serem juridicamente tuteladas que implicam não apenas transferência das medidas de controle conhecidas no mundo offline, mas também a criação de outras, reconfigurando o modo de produção do Direito, que passa a depender de uma ação mais integrada entre os atores da comunicação online e menos centralizada na figura do Estado.

Na tormentosa travessia da era analógica para a digital, do conhecimento transferido em livro para o intercâmbio online, das interações presenciais para contatos virtuais, os boatos se

revestiram na dinâmica tecnológica e, ainda que não sejam fenômeno novo, as fake news ganharam uma dinâmica global cujos efeitos envolvem a discussão sobre os limites da liberdade de expressão online e os respectivos modelos de modulação e moderação de conteúdo.

Estar conectado não é mais uma opção, quase todas as transações modernas atravessam caminhos virtuais e, por isso, as pessoas estão cada vez mais dependentes da internet. Desse modo, é preciso buscar medidas de contingenciamento para as fake news, ainda que não sejam definitivas nem suficientes para a preservação e proteção de direitos online, porque a reformulação do quadro geral, com o fim dos monopólios e mais competitividade no mercado de modo a alterar o modelo de negócios online, ainda parece longínqua. Na atual arquitetura da comunicação digital, os gestores do fluxo comunicacional, apesar de serem empresas privadas, tem mais poder que muitas jurisdições. De forma precisa, Silveira (2018, p. 37) trata sobre a questão:

As plataformas reúnem pessoas que querem ou necessitam se agrupar ou pertencer a redes de amizades, negócios, afetos, entretenimento. Como integrantes, essas pessoas têm o poder de entrar ou abandonar a plataforma, muito diferente do poder que os gestores ou que os donos dessas redes privadas detêm. Um dos principais modos de controle que os gestores das plataformas possuem sobre seus usuários se dá pela modulação das opções e dos caminhos de interação e de acesso aos conteúdos publicados. (SILVEIRA, 2018, p. 37)

Diante disso, a governança digital passa a ser uma preocupação mundial, no sentido de que o exercício de poder pelas empresas gestoras dos fluxos comunicacionais seja limitado e respeite direitos fundamentais. O desafio é submeter as empresas gestoras dos fluxos comunicacionais online a um modelo de responsabilização mais forte para resguardar direitos fundamentais e preservar instituições como a própria democracia. Ocorre que se expandiram sob um sistema regras frouxas, ascenderam como grandes potências privadas e têm forte influência econômica, por isso a concretização de direitos em seus espaços não depende apenas de medidas legislativas estatais, mas também de uma abordagem do modelo de gestão dos fluxos que considere a arquitetura do design em rede.

Nesse sentido, Evelyn Douek (2022) argumenta ainda que a imagem da moderação de conteúdo padrão, individualizada, é instantânea no tempo e por isso estática, desatualizada e contrastante com a dinâmica online que é fluida tendo em vista que as próprias empresas se reinventam, criando novas possibilidades para reagir e compartilhar conteúdo, além de atualizações quanto as frequentes mudanças de significado de expressões que exigem contínua revisão da moderação de conteúdo, a exemplo dos significados de hashtag e de memes.

A autora (DOUEK, 2022) também afirma que os designers de sistemas de moderação fazem escolhas e a resposta certa sobre uma moderação de conteúdo varia de país para país, de acordo com suas leis e com a concepção local da liberdade de expressão, por isso, em situações de risco, preferem errar mais e excluir mais conteúdo falso-positivo do que arcar com o custos de falsos-negativos. Diante disso, ela argumenta que os erros de moderação em escala são inevitáveis tanto na exclusão em excesso como também nas falhas em não retirar conteúdo por avaliá-los incorretamente. Além do modelo de moderação de conteúdo caso a caso, a autora traz outras formas de controle do discurso que não são abordadas ao se priorizar o paradigma offline.

A moderação baseada no comportamento de grupos de contas e dos atores ou associações por trás deles tem proporções muito maiores e mais opacas que a moderação de conteúdo individual. O foco são as condutas ilegítimas que agem de forma coordenada e que isoladamente podem não representar perigo. Segundo a autora, a moderação comportamental é realizada por equipes relacionadas à segurança cibernética e se opera com a colaboração das empresas entre si e com governos que compartilham informações e suas decisões. A atuação das milícias digitais pode se encaixar nessas situações de atividade coordenada em rede em que as remoções de contas e de conteúdos são baseadas em padrões de comportamentos.

Segundo a autora (DOUEK, 2022), nas situações em que as empresas tomam decisões conjuntas sobre moderação de conteúdo comportamental com remoção de conteúdo ou de perfis, os usuários podem nem tomar conhecimento de que foi uma decisão coletiva e, mesmo que tenham ciência dessa circunstância, o remédio será uma reclamação individual em cada empresa ou contra cada uma em âmbito judicial. São obscuros os critérios utilizados para definir o comportamento inautêntico e os gestores da comunicação justificam a falta de transparência como medida necessária para impedir que os infratores conheçam as formas de detecção e escapem do controle.

Ainda de acordo com Evelyn Douek (2022), existe a moderação tomada por decisão delegada em que terceiros avaliam o conteúdo, verificam os fatos e julgam a falsidade ou não da postagem, sendo uma forma das empresas transferirem a responsabilidade por erros de moderação e aparentarem neutralidade. Segundo a autora, esse foi um mecanismo muito usado na pandemia de Covid-19 em que a OMS – Organização Mundial de Saúde avaliava o quê era desinformação ou não no sentido de manter a população mais bem informada.

Outra modalidade de moderação mencionada pela autora (DOUEK, 2022) e que escapa a concepção tradicional de controle de conteúdo caso a caso é a moderação baseada no design

e nos recursos tecnológicos em que há possibilidade de controle de conteúdo antes mesmo da publicação como perguntas e advertências sobre o conteúdo dirigidas aos emissores bem como a existência de bancos de imagens pré-catalogadas como de pornografia infantil, por exemplo, que impedem o download.

Evelyn Douek (2022) menciona também a possibilidade de moderação de conteúdo com base no design como ocorre na redução da frequência em que um assunto é visualizado nos feeds, tornando ainda mais difícil para o usuário identificar o controle. Para a autora, dar maior autonomia ao usuário sobre a amplitude de suas postagens com a opção de restringir o acesso da publicação para amigos selecionados, ou para escolher quais assuntos tem interesse, excluir comentários sobre suas postagens dentre outras é também moderação de conteúdo com base no design.

Em seguida, ela (DOUEK, 2022) enumera as razões que justificam suas concepções e que rejeitam a moderação individual como modelo padrão. Em primeiro, os casos individuais não são um bom paradigma para evitar e corrigir erros sistêmicos e escolhas do design ex ante porque as nuances de um caso podem não refletir em outros, depois porque corrigir erros individuais ex post se mostra ineficaz para reforma de sistemas que não são, nem formam, uma unidade, tendo em vista que cada empresa tem suas regras internas. Em terceiro, a falsa transparência de um modelo individual causa o embaçamento sobre a percepção da moderação sistêmica como a automatizada. Por último, ela afirma que a maximização de regras processuais individuais obscurece as compensações do controle sistêmico em que há proteções individuais adicionais, como o contingenciamento em tempo real e aquele que impede a violação de direitos autorais, o download de imagens de pornografia infantil pré-catalogadas e o que adverte o emissor diante de um conteúdo ofensivo.

Ocorre que, em muitas jurisdições é conferida imunidade de responsabilidade aos gestores e mediadores da comunicação online e não lhes incumbe o dever de monitoramento prévio do que circula em seus domínios e isso diminui a incidência de formas tecnológicas de supervisão dos fluxos comunicacionais, então, as fake news circulam mais livremente.

Sob inspiração americana, o Brasil, acompanhou a posição preferencial da liberdade de expressão, tendo o STF decidido na ADPF nº 130, julgada em 2009, a preferência *prima facie* da liberdade de expressão quando em choque com outros direitos fundamentais. Desse modo, o sistema jurídico brasileiro tanto em relação ao poder público quanto em demandas que envolvem atores privados e cidadãos em face da liberdade de imprensa, como liberdade comunicativa, adota uma proteção de primazia da liberdade de expressão.

Felipe Soares e Rafael Mansur (2020, p. 40) colocam em xeque a posição de primazia *prima facie* da liberdade de expressão quando em conflito com direitos da personalidade, porque a Constituição Federal não autorizou a preponderância da liberdade de expressão. Eles enfatizam que, na ponderação, a circunstância de preferência é uma escolha de índole subjetiva do julgador e não serve como diretriz hermenêutica. No mesmo sentido, Mart Susi e Robert Alexy (2020, p. 27) sustentam que o elemento ontônico da ponderação é o igual merecimento entre os direitos em conflito, razão pela qual não se poderia falar em preferência à liberdade expressão, sob pena de estabelecer um desequilíbrio antes mesmo da ponderação.

Antes do Marco Civil da Internet (MCI), Lei 12.965/2014, a construção jurisprudencial no Brasil, especialmente no STJ¹⁴, se manifestava pela possibilidade de responsabilidade das empresas gestoras dos fluxos por postagens de terceiros desde que, tendo sido avisadas sobre o conteúdo, não tomassem providências, ainda que o aviso não tenha sido feito através de ordem judicial específica. Tem-se então que o entendimento jurisprudencial do STJ¹⁵ antes do MCI é indicativo de que não há impedimento constitucional para o controle feito pelas empresas de gestoras dos fluxos comunicacionais. No entanto, após a vigência do MCI, a obrigatoriedade de exclusão de conteúdo passou a incidir somente quando a notificação for mediante ordem judicial específica.

Assim, o MCI no art. 19, em sintonia com a legislação e jurisprudência americanas confere imunidade às empresas de tecnologia que fazem a gestão dos fluxos comunicacionais. De acordo com esse dispositivo legal, as situações de tutela preventiva para limitar o discurso são excetivas e a decisão judicial sobre o legítimo exercício da liberdade de expressão é caso a caso e não define em termos gerais o que pode ou não ser dito, ou seja, nos termos do art. 19, o modelo de controle vigente no Brasil por publicações de terceiros eximiu as redes do dever de monitoramento prévio sob o receio de promover censura.

A adoção dessa medida visa combater a indústria de notificações para remoção de conteúdo. O Marco Civil assume posição de defesa da liberdade de expressão e garante aos provedores a necessária imunidade que minimiza o temor que poderia existir no sentido de que a não remoção do conteúdo, depois da notificação geraria a sua responsabilização. Caso contrário, ao se determinar a remoção de conteúdo depois do envio de uma simples notificação, conteúdo lícito poderia ser removido, tão só e unicamente, para que o provedor não se responsabilizasse. (GIACCETTA, 2020, p. 289)

¹⁴ Ver mais em STJ, REsp 1642997 / RJ, Recurso Especial 2016/0272263-4, Relatora Ministra Nancy Andrighi, T3 - Terceira Turma. Julgado em 12/09/2017. Publicado no DJe em 15/09/2017. Disponível no site <https://processo.stj.jus.br/SCON/pesquisar.jsp>.

Para Schreiber (2020), a opção legislativa brasileira visou impedir o denunciamento em rede que poderia levar as plataformas a excluir conteúdo para evitar a responsabilização provocando assim um efeito silenciador de minorias e grupos vulneráveis trazendo como consequência a diminuição da pluralidade de opiniões no ambiente online. Ocorre que, segundo ele, esse modelo cria uma dificuldade para a vítima ao estabelecer a necessidade de ordem judicial específica para excluir conteúdo em rede, com todos os custos e a demora das formalidades legais, enquanto os danos continuam a gerar efeitos.

Diante disso, tramita no Supremo Tribunal Federal (STF) (BRASIL, 2023a) o Incidente de Repercussão Geral nº 987 no Recurso Extraordinário 1.037.396/SP, relator Ministro Dias Toffoli, cujo tema é a discussão sobre a constitucionalidade do art. 19 da Lei n. 12.965/2014, Marco Civil da Internet (MCI) que determina a necessidade de prévia e específica ordem judicial de exclusão de conteúdo para a responsabilização civil de provedor de internet, websites e gestores de aplicativos de redes sociais por danos decorrentes de atos ilícitos praticados por terceiros. A transcendência foi justificada pelo efeito multiplicador da presente demanda bem como pela relevância que a internet assumiu na vida dos cidadãos, ao tornar-se parte imprescindível da sociedade atual.

A discussão trata sobre o dever de monitoramento das redes e a respectiva responsabilidade dos provedores de internet, websites e gestores de aplicativos por manifestações ilícitas dos usuários. Eventual declaração de inconstitucionalidade do art. 19 do Marco Civil da Internet pode determinar novo modelo de responsabilidade civil e, por consequência, novas formas de controle de conteúdo.

O caso que motivou a demanda foi o pedido da autora para exclusão de perfil falso feito com seus dados e indenização com fundamento na responsabilidade objetiva do art. 17 do CDC. Desde a propositura da ação, muitos anos se passaram e atualmente algumas redes já contam com mecanismos para exclusão de perfis falsos e canal de denúncias, ainda que não sejam obrigatórios. No entanto, persiste a polêmica sobre a responsabilidade a ser definida, pelo Código de Defesa do Consumidor (CDC) ou Marco Civil da Internet (MCI). Ambas as leis têm mesma hierarquia, sendo a Lei 12.965/2014, conhecida como Marco Civil da Internet, posterior e específica quanto ao regramento do espaço virtual.

Para Balkin (2018), quando são atribuídos aos usuários o controle do conteúdo online, por meio de denúncias e reclamações, eles passam a trabalhar para as empresas, fornecendo informações que as ajudam a impor os padrões da comunidade. Como essas empresas não são transparentes, quando necessário ou conveniente, dispensam as políticas de governança e a

moderação de conteúdo se torna arbitrária. Assim, sob a perspectiva das empresas gestoras dos fluxos comunicacionais, o modelo brasileiro é interessante e menos custoso já que a moderação se opera a partir dos usuários e não de moderadores próprios ou sistemas de filtragens internas.

Evgeny Morozov (2020) percebe o problema de se relacionar o intercâmbio de informação à responsabilidade do usuário, transferindo para o indivíduo o combate das fake news, como ocorre numa relação de consumo, por meio das denúncias dos usuários. A relação entre cada usuário e a empresas de tecnologia pode ser qualificada como um contrato segundo as regras do direito civil ou de acordo a legislação consumerista, no entanto, a circulação da informação online não deve ser tratada segundo a lógica financeira porque impacta na formação cidadã e resvala na democracia.

O Brasil, ao eleger a via judicial com ritos não propriamente céleres para enfrentar uma dinâmica veloz como a divulgação de fake news em ambiente online, coloca as empresas de tecnologia uma posição cômoda, em que não lhes são atribuídos o dever de monitoramento do conteúdo em rede nem a proteção da vítima. Diante disso, Ingo Sarlet e Ivar Hartmann (2019) criticam a circunstância do Poder Judiciário continuar a centralizar o controle de conteúdo online em razão da quantidade de manifestações diárias no ambiente online.

A Alemanha, contrariamente ao Brasil, implementou um modelo de moderação de conteúdo em que impõem obrigações às empresas gestoras dos fluxos comunicacionais. Ao editar a NetzDG o país institucionalizou a obrigatoriedade das gestoras do fluxo de comunicação online, com mais de 2 milhões de usuários domésticos (artigo 1, § 1, 1 e 2), estabelecerem um canal direto, online e constantemente acessível para reclamação (artigo 1, § 1 3 da NetzDG). Desse modo, a NetzDG regulamentou o procedimento da reclamação direta dos usuários em que apontam e especificam postagens ilícitas, denunciam o conteúdo e pedem a retirada da mensagem ofensiva (artigo I §3 1 e 2 da NetzDG).

O conteúdo ilícito deve ser de fácil reconhecimento e está inserido dentro das tipificações criminais enumeradas na lei (artigo 1 § 1, 3 NetzDG). Além disso, a NetzDG determinou que as reclamações dos usuários tenham uma decisão, ou seja, que o pedido para remoção de conteúdo feito pelos usuários seja analisado em atenção às regras de transparências, com solução proferida por um corpo destacado de profissionais exclusivamente para o fim de julgamentos dentro de um prazo fixado em lei.

A NetzDG determina também a obrigatoriedade de relatórios semestrais para monitoramento e fiscalização do mandamento legal (artigo 1, §§ 1 e 2 da NetzDG) e a aplicação

de multas vultuosas em caso de reiterado descumprimento (artigo 1,§ 4 da NetzDG). As sanções pecuniárias impostas às plataformas decorrem do descumprimento legal de prestar informações detalhadas ou sobre falhas reiteradas no processamento das reclamações, portanto, não são aplicadas como medidas reprovativas de uma decisão destacada.

O ineditismo da NetzDG custou severas críticas em virtude das restrições à liberdade de expressão online, atacando-lhe inclusive de inconstitucionalidade. Ricardo Villas Boas Cueva (2020) elenca as principais críticas a NetzDG, como a de que as pesadas multas ocasionariam bloqueios e exclusões de conteúdo aleatórios comprometendo o modelo de negócios e a liberdade de expressão online, alegação de que a NetzDG não auxiliaria na identificação e responsabilização do criminoso que disseminou discurso de ódio ou fake news, e argumentos no sentido de que a referida lei não tinha por pressuposto o apagamento de conteúdo ilícito mas apenas estabelecer rotinas empresarias para que as manifestações não se perpetuassem.

Para Balkin (2018), a NetzDG é um modelo de censura colateral tendo em vista que os Estados forçam as empresas da infraestrutura digital a bloquear e excluir conteúdo dos usuários sob pena de pesadas multas. Diante disso, se torna economicamente mais interessante restringir conteúdo ao vez de mantê-lo online. O autor também afirma que esse modelo é um acordo entre o Estado e as empresas que implementam políticas de discurso mediante uma burocracia privatizada.

Para Evelyn Douek (2022), os relatórios de transparência sobre o processamento das denúncias não podem ser analisados friamente por que a moderação baseada apenas em números de exclusões não avalia o mérito dessas decisões e pode estimular remoções excessivas. A autora argumenta ainda que os relatórios fornecem um retrato estático que não considera outras formas de moderação de conteúdo que acontecem online como as pressões governamentais ou desacordos com normas internas e diretrizes da comunidade.

O Instagram faz uma referência a NetzDG (INSTAGRAM, 2022c) e explica que procede a abordagem em duas etapas para atender ao mandamento da referida lei, tal circunstância demonstra como a disciplina do ambiente virtual é desafiadora tendo em vista que as possibilidades técnicas muitas vezes encontram caminhos que beiram a desobediência. Assim, o Instagram só exclui definitivamente de seus domínios conteúdo denunciado que contrarie as “Diretrizes da Comunidade”. Caso o conteúdo denunciado não contrarie as Diretrizes da Comunidade mas infrinja a NetzDG o acesso será impedido apenas na Alemanha, permanecendo online e disponível em outros países.

Note-se que o modelo da Alemanha, de autorregulação regulada, repete, de certa forma, as conhecidas medidas de moderação de conteúdo que se baseiam em análises caso a caso que é um modelo custoso e tende a se tornar cada vez mais lento, tendo em vista o volume de reclamações, e não menos falível do que modelo judicial porque os julgamentos humanos, mesmo sendo moderadores altamente treinados, estão sujeitos a subjetivismos, influências culturais, religiosas, ideológicas dentre outras.

Diante disso, Evelyn Douek (2022), afirma que a imagem padrão de moderação de conteúdo online tem se concretizado numa abordagem caso a caso, semelhante ao sistema de revisão judicial tradicional, em que devem ser asseguradas a comunicação sobre da medida de moderação, com razões fundamentadas individualmente, e a oportunidade a defesa para os usuários inclusive com direito de apelação e revisor humano.

No entanto, de acordo com a autora, a moderação de conteúdo caso a caso, feita de acordo com apelos individuais foca nas informações prestadas pelos tomadores de decisão sobre como resolvem os casos e ignora outros tipos de moderação como a influência dos anunciantes, a solicitação governamental para remoção de conteúdo segundo demandas formais legais ou mecanismos informais, expectativas dos usuários e não considera novas formas de intervenção futura, em consonância com os avanços tecnológicos, por isso a autora defende a necessidade de avaliar o devido processo sistematicamente considerando a diversidade e o dinamismo dos ordenamentos privados.

Diante disso, a autora (DOUEK, 2022) conclui que, se continuar nesses moldes, o modelo se mostrará impraticável, por que a revisão humana é cara e se tornará cada vez mais demorada em virtude de reclamações em escala, de modo que a maioria das denúncias não receberá resposta porque a capacidade de resolução é inferior à demanda. Ela argumenta ainda que exigências processuais excessivamente onerosas podem incentivar as empresas a proibir menos discurso com a finalidade de lhes diminuir os encargos de fiscalização ou tornar mais difícil o uso dos mecanismos de denúncia para reduzir o número de reclamações.

Além do volume em massa, os conflitos no ambiente digital assumem novas modalidades de organização social com repercussões inéditas quanto ao tempo, espaço e velocidade em virtude das dimensões virtuais e que tornam obsoleto o padrão de moderação de conteúdo caso a caso. Por isso, um modelo de moderação estático baseado no julgamento de casos individuais ex post e em decisões binárias de remoção ou desistência de exclusão de conteúdo, tal como a revisão judicial, não atende a necessidade de uma abordagem sistêmica

que considera o dinamismo do ambiente digital com formas de controle abrangente, automatizado e em tempo real.

3.3 Milícias digitais

A comunicação online impactou em muitos aspectos as dinâmicas sociais e possibilitou mobilizações das pessoas através da internet que se agruparam por identidade de interesses e outras afinidades, políticas ou não. Os usuários perceberam o potencial desses movimentos como poder social de reivindicação e transformação, a exemplo da Primavera Árabe em que protestos públicos eram organizados através das redes, reunindo multidões em pouco tempo e com baixo custo de divulgação (MAGRANI, 2014).

Também são expressões desses movimentos articulados em rede o denominado shaming que é o ranqueamento de empresas pelos consumidores, que se associam para direcionar as insatisfações e protestos contra uma marca ou um produto, como uma espécie de sanção social pela reprovação de suas condutas. Para Schreiber (2020), o shaming e o cyberbullying configuram exercício abusivo da liberdade de expressão porque a sanção social é imposta de modo sumário, sem chance de defesa. Nesse mesmo sentido, Byung Chul Han (2019) menciona os shitstorms que não questionam o poder estabelecido apenas embaraçam e escandalizam pessoas individuais como candidatos.

A articulação em rede, apesar de ser um grande benefício tendo em vista a ampliação da redes de contatos, também possibilitou ondas de imigração, favoreceu o terrorismo que arregimenta membros e engendra ataques pela internet, disseminação de fake news através de milícias digitais além de outras atividades criminosas que também se valem dessa estrutura interligada.

Pierre Levy (2018, p. 129) afirma que “a conexão é um bem em si” de modo que o sistema internacional da internet é uma complexidade de redes que se interconectam podendo os usuários usarem a estrutura comunicativa para diversos fins. Desse modo, o ativismo em série pode promover uma celebridade a nível mundial como ocorreu na situação em que a música da artista brasileira Anitta foi incluída em diversas playlists e a reprodução repetidas vezes a fez atingir o topo do ranking global de serviços de streaming (VEJA, 2022). A Reportagem da BBC (2017) refere-se aos ativistas em série como pessoas reais que são

extremamente ativas em rede. Nessas situações o movimento articulado em rede destruiu a realidade compartilhada e criou artificialmente outra que beneficiou a cantora.

Além dos ativistas online em série, existe um exército de robôs que atua no mundo digital e que participa intensamente no compartilhamento e produção de conteúdo cujos perfis podem ser integralmente operados por máquinas ou coordenados por humanos, esses últimos denominados de trolls ou ciborgues. São formas diferentes de interação em que as trocas de mensagens não agregam valor nem estabelecem diálogo porque não há introspecção ou reflexão, mas mera reprodução automática.

Nesse sentido Wilson Ceron (2022) esclarece que os bots¹⁶ podem ser usados para fins maliciosos porque ao promoverem uma frequência de publicação elevada acabam por promover um maior alcance do conteúdo, destacando determinado assunto ou discussão ou posição como relevante, afetando, desse modo, a percepção da realidade, as prioridades da pauta de discussão pública e o resultado das decisões políticas.

Diante disso, é importante que o usuário saiba com quem interage, na verdade, com o quê, porque sendo máquinas, a atividade de demonstrar que um conteúdo compartilhado se trata de fake news é ineficaz, por exemplo. Responder e comentar gera mais visualização e mais lucro a quem criou ou divulgou fake news tendo em vista que as mensagens são impulsionadas e remuneradas de acordo com a reação dos usuários. Portanto, nas situações de perfis artificiais, mesmo que seja híbrido, administrado em parte por humanos e em parte por robôs, a estratégia de enfrentamento deve ser tecnológica porque nessas situações a racionalidade humana é substituída pelo tecnicismo automatizado.

Ocorre que, como dito na reportagem referida acima (BBC, 2017), identificar robôs e ciborgues demanda tempo e dinheiro, mas como ponto de partida relaciona alguns padrões e atividades que podem ser usados como indicativos de perfis não-humanos, quais sejam: número de vezes que replicam um conteúdo, desproporção entre seguidores e o número dos perfis que seguem, pobreza vocabular e palavras repetidas, quantidade de menções a outros usuários, datas coincidentes de criação, ativação e desativação da conta.

O uso de robôs e ciborgues para propagar e impulsionar desinformação no ambiente online é amplamente conhecido, nesse sentido Vânia Baldi (2018) ressalta que a atividade de trolls e bots é semear discórdia e hostilidade entre os grupos que rivalizam pontos de vista

¹⁶ Para melhor esclarecimento, utiliza-se Bot, que é a abreviação de robot, sinônimo de robô em inglês.

opostos. Para a autora, apesar dos mecanismos tecnológicos não serem a razão pelo quadro social fragmentado e polarizado, as práticas online induzem a radicalização das opiniões e uma aceleração do tempo social, ou seja, uma cultura não reflexiva sobre as informações com a formação de opiniões fast food de modo que o ineditismo, e não a verdade, é que movimenta o ciclo de compartilhamento em rede.

Segundo Callejón (2022a) as fake news não tem o interesse na democracia ou numa formulação política alternativa ao regime, desse modo, para o autor, as fake news promovem uma destruição da realidade, transformando a liberdade de expressão num produto comercial e deixam de servir à democracia e ao pluralismo político. A divisão radical no ambiente online ocasionou o que ele denomina de pluralismo de posições monistas em que visões de mundo diferentes não reconhecem as demais, estabelecendo uma incomunicação cujo efeito é potencializado pelo uso de algoritmos. Desse modo, a atuação das milícias, com intensificação da disseminação viral de fake news, é ainda mais destrutiva para os processos comunicativos.

Ocorre que se as fake news pretendem a destruição da realidade compartilhada, sem interesse específico, os movimentos articulados em rede, como as milícias digitais, têm objetivos mais definidos. No caso da cantora Anitta, houve uma ação com o objetivo de promover uma música e uma artista, ou seja, se o modelo de negócios não tem um interesse específico para privilegiar uma opção política, a falta de controle do comportamento miliciano, enquanto atividade coordenada online, possibilita a criação artificial de cenários, como a eleição de um determinado candidato.

A realidade social dividida, racionalmente precipitada e impulsiva, facilmente suscetível a conteúdos emocionais e manipulável pela lógica algorítmica, que modula as informações segundo um viés confirmatório de cada perfil, tem favorecido a destruição da realidade. O disparo sistemático de desinformação segundo uma tendência específica, não apenas prejudica a leitura dos fatos, mas condiciona e direciona a percepção das pessoas e influencia ou mesmo determina suas escolhas. Por essas razões, as práticas online têm sido objeto de regulação.

No Brasil, o projeto de lei 2630 de 2020 que ficou conhecido como lei das fake news e atualmente se propõe a ser a Lei Brasileira de Liberdade, Responsabilidade e Transparência na Internet visa regulamentar redes sociais e serviços de mensageria privada que contam com mais de 2 milhões de usuários no país. Em vários dispositivos do projeto é visível a preocupação com a desinformação online e com as consequências cívico-democráticas.

No art. 3º do projeto de lei 2630 de 2020 são relacionados três objetivos da lei, sendo o do inciso III “desencorajar o uso de contas inautênticas para disseminar desinformação nas aplicações de internet”. Isso ocorre porque as guerras de narrativas entre grupos opostos, valendo-se de violência argumentativa e de encadeamento de argumentos baseados em fake news, se escondem por detrás de falsos perfis e acabam alimentar uma sensação de que posições radicais e extremistas são majoritárias. Numa interminável troca de mensagens de conteúdo falso, enganoso e descontextualizado os usuários se envolvem em discussões intermináveis sem qualquer benefício para o diálogo democrático e para o debate político, criando o medo do isolamento social em quem diverge desses perfis.

Nesse sentido, o referido projeto de lei traz também dispositivos para disciplinar as atividades online relacionadas a desinformação e ao uso de contas inautênticas e disseminadores artificiais, entendidas como perfis falsos ou automatizados. O artigo 4º, inciso IV, define contas inautênticas como a “conta criada ou usada com o propósito de disseminar desinformação ou assumir identidade de terceira pessoa para enganar o público;”. No inciso seguinte, V, do mesmo artigo conceitua “disseminadores artificiais” como sendo “qualquer programa de computador ou tecnologia empregada para simular, substituir ou facilitar atividades de humanos na disseminação de conteúdo em aplicações de internet.”

A democracia se constitui num regime de poder em que o regramento jurídico é aprovado após deliberação popular, diretamente ou através de representantes, em que o processo de discussão e votação é previamente conhecido, num rito que avança em fases ordenadas e segundo prazos, onde há oportunidade para manifestação de posicionamentos divergentes no intuito de que as contribuições possam influir na decisão final. Desse modo, o resultado do processo legislativo tende a ser uma construção compartilhada.

Ocorre que no ambiente de discussão online essa interação dialógica entre os usuários resta prejudicada porque os algoritmos atuam para reforçar os pontos de vistas iniciais ao direcionar conteúdo segundo o viés de confirmação de cada um, distanciando os que se opõem. Além disso, não há procedimento estabelecido para tomada de decisão, como as regras procedimentais do jogo democrático, defendidas por Bobbio (2019).

Desse modo, a distribuição informativa do ambiente online não é balanceada e por isso não favorece ao consenso democrático, pelo contrário, visa tumultuar a discussão na medida em que sobreposições de conteúdos que confirmam concepções promovem a agitação dos discursos online e se compatibilizam com o lucro. O comportamento de disseminar fake news e desconstruir artificialmente os fatos enquadra-se como uso malicioso dos mecanismos online,

numa nítida desconfiguração dos fins originários da internet, usurpados pelos interesses econômicos.

A sistemática de bombardear fake news e desencadear a participação online contribui para uma percepção distorcida da realidade natural e tem trazido implicações graves para a democracia. Segundo Vânia Baldi (2018, p. 05), camadas infindáveis de versões e opiniões criam um ânimo geral de desconfiança, conferindo primazia às interpretações em relação aos próprios fatos, desprezando argumentos lógicos e referências científicas e histórias.

Apesar da manipulação dos fatos ser conhecida tática política sintetizada na frase de Joseph Goebbels, ministro da propaganda nazista, segundo a qual “Uma mentira dita mil vezes torna-se verdade.”, as atividades milicianas representam perigo porque não apenas destroem a realidade compartilhada, mas constroem cenários fictícios para influir na realidade pragmática, supervalorizando o falso e as versões em detrimento da verdade e dos fatos. No comportamento miliciano há um aprimoramento da atividade de produzir e distribuir fake news, não meramente determinada pela lógica de remuneração pelo engajamento.

Por isso, o projeto de lei 2630/2020, no art. 14, determina que o uso de disseminadores artificiais e administradores de contas deve ser comunicado ao provedor de mensageria privada sob pena de exclusão da conta em razão da movimentação e número de postagens incompatível com o uso humano. O uso de disseminadores artificiais confere maior amplitude ao usuário tendo em vista que sua voz passa a ter uma audiência maior, desse modo quem pode dispor desses mecanismos é privilegiado porque passa a atingir um espectro maior de pessoas em comparação a ressonância normal de um perfil humano.

As campanhas de viralização digital, como a das milícias digitais que disseminam fake news, com o uso de disseminadores artificiais e de contas inautênticas tem o potencial para deturpar juízos de valor das pessoas em relação aos acontecimentos ou mesmo paralisar o pensar diante da quantidade de informação despejada, tornando automatizada a vontade dos usuários/cidadãos. Os padrões e comportamentos algorítmicos rivalizam com a racionalidade humana e, por se apresentarem como uma atividade incessante de semear dúvida ou incerteza, são poderosos mecanismos que embaraçam o pensamento e manipulam dos fatos.

As práticas digitais inauguram novas relações. Para Byung Chul Han (2019) o enxame digital se caracteriza como uma aglomeração de indivíduos singularizados, individualmente identificáveis em seus perfis apesar de se apresentarem anonimamente. Para ele, o enxame não se constitui uma massa uniforme e nem desenvolve um sentimento de identidade, os objetivos

são efêmeros e rapidamente se dissolvem. Segundo o autor, no enxame digital há interações momentâneas que tem o poder de contestar e mudar, mas não de avançar em projetos políticos duradouros.

As milícias digitais, assim como o enxame, também interferem na esfera política e são manifestações dos novos arranjos online, mas enquanto que o enxame é uma reação social, como que uma convulsão que rapidamente se aglomera e se dissipa, faz evoluções e não tem uma direção específica, as milícias digitais que disseminam fake news com o uso de robôs se valem do poder econômico ao dispor dessa tecnologia para atingir um objetivo definido, segundo uma finalidade fraudulenta. A expressão “milícias digitais” é corriqueiramente tratada como vínculos tecnológicos negativos, com finalidades ilícitas. Por agregar muitos perfis, as milícias digitais se apresentam como uma ação coordenada em que podem estar associadas a robôs para potencializar os disparos maciços de mensagens. Nesse sentido Edilene Lobo Morais, José Luis Bolzan e David Nemer conceituam milícias digitais como:

[...] uma associação de pessoas interligadas de forma mais ou menos flexível e sem um arranjo jurídico-legal, que agem de maneira coordenada ou orquestrada na web, em sua grande maioria pelas redes sociais, se utilizando de robôs, contas automatizadas e perfis falsos, promovendo campanhas de ataques e/ou cancelamento de imagens e reputações de adversários ocasionais. (LOBO; MORAIS; NEMER, 2020, p. 260)

Como se percebe, o conceito acima se assemelha a comportamentos que se aproximam do shaming e do shitstorm na medida em que afetam a reputação das pessoas, celebridades, anônimos ou candidatos a cargos eletivos, ferindo-lhes direitos da personalidade. As milícias digitais agem sem limites de fronteiras e não são um fenômeno local, podem ter ramificações internacionais o que torna ainda mais difícil a contenção de suas atividades com medidas analógicas, expedidas fora do ambiente digital ou por aqueles que não detém os códigos-fonte para estancar o derrame online.

Numa sociedade em que os fatos se tornaram estatísticas e os acontecimentos foram convertidos em expressões matemáticas, se reconhece o valor de um corpo coletivo engajado como um bem valioso, capaz de manipular a realidade usando a lógica algorítmica. Nesse sentido, movimentos em rede, não personificados e informais, exercem influência social e geram efeito em cadeia ao alterar a percepção da realidade que é retratada em número de likes, compartilhamentos e seguidores. A quantificação ascende como uma característica da sociedade conectada e os números geram reputação para um perfil.

Desse modo, o uso de disseminadores artificiais e contas inautênticas com o disparo maciço de mensagens faz crer que o ponto de vista de um perfil, que se vale dessas estratégias

tecnológicas, é majoritário e essa percepção enganosa da realidade influencia e direciona a opinião pública e, assim os algoritmos, que destroem a realidade compartilhada, moldam falsas impressões.

Para Elisabeth Noelle-Neumann (1993), a opinião pública parte de duas concepções diferentes, a manifesta e a latente. A opinião pública manifesta tem o propósito democrático, de debate racional, de troca de ideias e opiniões para tomada de posição entre cidadãos informados. A opinião pública latente tem um elemento de coesão, de poder da opinião pública sobre um indivíduo pressionando-o em razão do medo do isolamento social.

Segundo a teoria do espiral do silêncio (NOELLE-NEUMANN, 1993), desenvolvida pela autora, os meios de comunicação de massa são fundamentais para as conclusões pessoais, através deles é possível observar o ambiente e avaliar quais as opiniões encontram aprovação na sociedade e quais levam ao isolamento. Para ela, as pessoas se comportam diferente diante das impressões de opiniões que julgam majoritárias, nesse sentido argumenta que há uma tendência para as pessoas aderirem às posições que parecem mais fortes e aceitas na sociedade. Assim como um espiral, acionam um processo silencioso que passa a atrair as pessoas e se robustece até que se consolidar e dominar a opinião pública.

Em tempos de aceleração social e da imediatez das massas frente às notícias, a atuação de milícias na propagação viral de fake news tem efeito fulminante na opinião pública, que reage instantaneamente do modo emocional às mensagens, fazendo pressão sobre o indivíduo/cidadão e causando enorme volatilidade do eleitorado. Ricardo Vilas Boas Cueva (2020) argumenta que, em relação às fake news, a verificação dos fatos é laboriosa e isso causa frustração no ambiente político gerando um sentimento de vertigem política. Assim, as fake news associadas a disseminação por milícias causam ainda mais instabilidade no comportamento dos cidadãos, atuando de modo sombrio sobre as dinâmicas sociais.

Também ao discorrer sobre o controle inconsciente do comportamento, Wolfgang Hoffmann-Riem (2021, p. 65) revela que os bots, “atores técnicos controlados por computadores em redes sociais que realizam tarefas repetitivas automaticamente e sem a necessidade de interação com comunicadores humanos”, podem destruir as comunicações de outros em razão das “contribuições provocativas ou deliberadamente falsas”. Desse modo, o ataque e o uso deliberado de fake news são formas abusivas da liberdade de expressão e podem ocasionar o silenciamento de posições divergentes e das minorias.

Por isso o autor defende que a proteção à liberdade inclui também a garantia para o uso da liberdade, ou seja, a liberdade de expressão não resta garantida quando lhe são impostos empecilhos para manifestação. Assim, se a expressão de um provoca um efeito silenciador em vozes divergentes, o comportamento e a liberdade de quem cala restam comprometidos e por isso se mostra importante conter quem silencia outras vozes.

Os processos comunicativos online e suas práticas, com novos arranjos sociais, impõem a necessidade de compreensão da internet como se apresenta no mundo fático em que os movimentos em rede, as milícias e uso de mecanismos tecnológicos potencializam vozes de alguns perfis ao conferir maior amplitude. No ambiente online, a voz não tem o mesmo valor para todos e nem sempre isso está cristalino por que, apesar da internet ser uma enorme abertura na mídia para os excluídos, o acesso aberto não é sinônimo de igualdade.

A correspondência entre real e virtual é enganosa. A igualdade na democracia se dá porque todos os votos têm o mesmo valor, mas isso não encontra correspondência na internet porque a voz não tem a mesma paridade para todos, mesmo que se julgue o acesso aberto como direito de voz no ciberespaço porque alguns se beneficiam de disseminadores artificiais e robôs e têm mais visibilidade. A voz no ambiente online não equivale ao voto numa democracia, mesmo que haja uma enorme similitude entre voz e voto no mundo real.

O poder de voz de um usuário pode ser amplificado artificialmente com disseminadores artificiais e perfis falsos, desigualando a relação entre os usuários, e isso só reforça as diferenças entre esfera pública e ambiente de discussão online. As milícias digitais são um fator que desequilibra, que traz mais em visibilidade a um conteúdo em virtude da disseminação com intenso disparo mediante comportamento coordenado entre muitos perfis, inundando as redes, fazendo parecer que a voz de poucos perfis sejam maioria e com um agravante: as impressões disseminadas de modo articulado e que destroem a realidade compartilhada têm potencial para criar falsos cenários.

Como abordado acima, no ciberespaço existem atores e mecanismos não visíveis que atuam disputando força e poder com repercussões para a liberdade de expressão. Reconhecer que o ciberespaço se trata de um modelo disruptivo, diferente do sistema analógico de comunicação e com modos de interferência no comportamento social ainda pouco conhecidos exigem um olhar mais apurado sobre as práticas das empresas de tecnologia, os movimentos em rede articulados inclusive o de milícias e que não se assemelham a modelos democráticos, no intuito de desenvolver formas efetivas de controle e responsabilidade dos gestores dos fluxos comunicacionais online.

3.4 Filtragem algorítmica

A moderação e modulação de conteúdo são, em regra, formas de filtragem algorítmica em que há uma seleção do fluxo comunicacional por algoritmos seja para excluir, bloquear, dimensionar visibilidade ou direcionar conteúdo. O controle de conteúdo online de modo automatizado teve inicialmente o objetivo de evitar que o ambiente online se tornasse um meio para violações de direitos autorais, facilitador para o terrorismo ou um canal para distribuição de pornografia infantil (NITRINI, 2021, p.53).

Ocorre que, à medida que a gama dos problemas e dos bens jurídicos a proteger se ampliaram, dilatam-se também os critérios de filtragem, operados de maneira opaca. Então, mesmo tendo sido uma iniciativa louvável, a implementação atual da moderação online beira a arbitrariedade porque não há definição clara do que seja proibido ou permitido, nem nos termos de serviço dos contratos, nas diretrizes da comunidade nem na legislação infraconstitucional.

Apesar da forma arbitrária que a moderação de conteúdo online tem sido operada pelas empresas de tecnologia, a verdade é que os usuários esperam que seja restringido muito mais do que conteúdo ilegal, então exclusão ou bloqueio de imagens de violência explícita, spam, pornografia, conteúdo extremista e que violem direitos da personalidade normalmente correspondem a anseios dos usuários.

A internet não corresponde mais aos ideais iniciais de democratizar os meios de comunicação para as minorias ao conferir-lhes voz porque apesar do acesso aberto nem todos tem a mesma oportunidade de serem ouvidos. A infraestrutura da comunicação online é um modelo de negócio baseado no lucro e, apesar da diversidade social com pluralidade de opiniões também ser importante, existe o perigo de que opiniões divergentes e dissidentes sejam filtradas, excluídas ou sombreadas do ambiente de discussão online aprofundando as desigualdades, privilegiando os que podem pagar por impulsionamentos de conteúdo e que terão mais visibilidade em suas postagens.

Callejón (2022a) entende que as empresas gestoras dos fluxos comunicacionais não têm um interesse específico para privilegiar uma opção política diante de outra e proceder varreduras de uma determinada visão político-partidária, suplantando o pluralismo. Ele argumenta, no entanto, que as empresas gestoras dos fluxos comunicacionais, não são neutras porque à medida que privilegiam fake news acabam por promover populismos que as

beneficiam indiretamente tendo em vista que os controles políticos sobre suas atividades empresariais se fragilizam.

Como modelo de negócios, essas empresas lucram com as guerras de narrativas entre posições rivais tendo em vista que a polarização e a radicalização mantêm a agitação de seus espaços e permitem captar a atenção da audiência por mais tempo, prolongando a navegação para extração de mais dados e para exposição dos usuários à publicidade. Desse modo, o receio de que as filtragens algorítmicas de fake news repercutam em restrições à liberdade de expressão a ponto de comprometer o pluralismo político online não parece compatível com o modelo de negócios porque o lucro decorre da pluralidade de seus espaços, da polarização. Para as empresas gestoras dos fluxos comunicacionais manter usuários que divergem, num ambiente plural, é mais interessante economicamente do que filtrar uma vertente política.

A matéria do *The Wall Street* (2020) revela que é possível tornar o ambiente de discussão online mais cívico e menos destrutivo com ajustes no design, no entanto, isso repercutiria em menor engajamento dos usuários, diminuindo os ganhos das empresas. Diante disso, as filtragens algorítmicas de fake news seriam prejudiciais aos modelos de negócios e não a liberdade de expressão, de informação ou ao pluralismo, como se costuma argumentar.

Para Loveluck (2018) a regulação da internet está sujeita a forte influência do design do hardware e software, e Balkin (2016, p. 370) complementa que “proteger a liberdade de expressão na era digital será cada vez menos um problema de discurso constitucional, embora essas proteções mantenham-se muito importantes, e cada vez mais um problema de tecnologia e regulamento administrativo.” Essas considerações conduzem a conclusão de que os controles na internet serão compartilhados entre agentes públicos e privados tendo em vista que o fluxo contínuo, a vigilância em massa e a necessidade de supervisão dinâmica dos espaços online necessitam de uma abordagem de moderação de conteúdo, *lato sensu*, em constante evolução.

De modo geral, a depender sob qual prisma se analisa e se estrutura as relações online, a regulação está inserida no dilema determinismo tecnológico versus construção social da tecnologia e das ideologias que os fundamentam. Disputam, no jogo de forças, normas técnicas versus instituições e regras jurídicas, sistemas com finalidades diversas. Como dito em tópico anterior no capítulo 2, item 2.2.2., as normas técnicas não são elaboradas por instâncias políticas tradicionais, no entanto fixam as práticas e prescrições em matérias como saúde, segurança e meio ambiente (FRYDMAN, 2018, p. 59). Diferentemente das regras jurídicas que passam pelo crivo de legitimidade democrática, uma vez que são deliberadas por órgãos de representação popular.

A corrente mais favorável aos interesses das empresas de tecnologia é o do determinismo tecnológico, termo criado por Thorstein Veblen e defendido por Andrew Feenberg, em que o poder político é obscurecido pela tecnologia, considerada como a maior força da sociedade moderna (FEENBERG, 2010). Por essa corrente, a autonomia das redes tende a prosperar e a *lex informatica* a ser livre, segundo a concepção liberal de mercado, havendo pouco espaço para regulação estatal, prevalecendo a autorregulação dos próprios serviços em detrimento da própria democracia.

Em contrapartida, caso se pretenda reestabelecer o controle governamental na relação entre usuários e empresas de tecnologia, a opção ideológica é a construção social da tecnologia. O constitucionalismo digital está inserido nesse contexto. As práticas das empresas de tecnologia estão sob a mira da lei e da atuação dos tribunais para que preservem direitos fundamentais dentro de seus domínios, ainda que as relações se estabeleçam entre particulares. Essa é uma visão mais afinada com as garantias processuais do Estado de Direito em que as empresas exploram economicamente o ambiente digital e não ficam blindadas de responsabilidades.

A corrente do determinismo tecnológico tenta convencer que os inventos são expressões do acaso e, para expandirem ao máximo suas potencialidades, requerem ampla liberdade, tendo em vista a imprevisibilidade dos avanços a serem descobertos (ZUBOFF, 2020, p. 557-558). A construção social da tecnologia pressupõe que os inventos decorrem da interpretação acerca das necessidades humanas num dado lugar e período, podendo-se concluir que as necessidades sociais impulsionam e direcionam as invenções (PINCH; BIJKER, 1984, p. 423).

Na competição entre tecnologia e sociedade, a corrente do determinismo tecnológico promove a supervalorização de normas técnicas, muitas vezes só compreensíveis por profissionais especializados, e que, por implicarem escolhas políticas e econômicas, acabam por determinar comportamentos (FRYDMAN, 2018). Considerando a influência recíproca entre ciberespaço e o mundo físico, a realidade pragmática de funcionamento das redes e as tensões em seu interior deve haver uma aproximação entre o determinismo tecnológico e a construção social da tecnologia, “um movimento quase que dialético entre a técnica e o social” (FREDES, MOLINARO, 2020, p. 258).

A circulação de imagens de violência, massacres religiosos com o de Christchurch na Nova Zelândia em 2019 ou suicídio de jovens ao vivo e ainda fake news que provocaram o linchamento de pessoas, além do discurso de ódio que ocasionaram genocídios (THE

GUARDIAN, 2018) foram situações que tornaram urgentes a necessidade de filtragens de conteúdo.

Pressões dos usuários, dos anunciantes, do Estado e econômicas motivaram as políticas de moderação porque nem usuários nem anunciantes querem estar associados a ambientes tóxicos e depreciativos. Evelyn Douek (2022) é categórica ao afirmar que se as plataformas que retirassem apenas conteúdo ilegal se tornariam inutilizáveis. Para a autora (DOUEK, 2022), toda moderação de conteúdo é de interesse do modelo de negócios e toda moderação de conteúdo implica em limitação da liberdade de expressão.

Nesse sentido Callejón (2022a) diz que no modelo de negócios da comunicação online, a liberdade de expressão nos contextos econômico e tecnológico sofre limitações, diferentemente no âmbito da política. Segundo o autor, no ambiente online não há narrativas que tragam risco as grandes corporações, aos rendimentos ou a capacidade de inovação. No entanto, as fake news proliferam na política, promovendo ataques e destruição da realidade, causando involução democrática.

Segundo o autor (2022), essa circunstância é um indicativo de que é possível preservar o tema 'política' como tem sido feito com a economia e tecnologia, não pelo exercício da liberdade de expressão, mas por configuração dos processos comunicativos realizados pelos gestores digitais. Assim, para Callejón (2022), a potencialização das fake news no ambiente de discussão online não está relacionada à ineficiência do controle algorítmico ou à incapacidade moral de diferenciar falso e verdadeiro, mas num interesse lucrativo em promover o falso.

Desse modo, a filtragem algorítmica de fake news disseminadas por milícias digitais pode ser um mecanismo sistêmico que não apenas contingencia fake news mas também controla tendências expansionistas dessas empresas que podem, ao permitir a destruição da realidade compartilhada e a construção de cenários artificiais, através de suas redes, pode estar sorrateiramente pondo em andamento o projeto de instrumentalização da humanidade, denunciado por Zuboff (2020).

Ainda que em regra, o controle editorial do ambiente online não seja como o da mídia tradicional que seleciona o conteúdo antes da publicação, as filtragens algorítmicas, os impulsionamentos ou mesmo os sombreamentos de conteúdo dimensionando a amplitude, a organização da informação trazida no feed, sequenciada com outras, modulada segundo preferências do perfil e acompanhada de comentários que se compatibilizam com a opinião do

usuário são estratégias que influenciam a pauta de discussão e o comportamento dos usuários (WIELSCH, 2020, p. 97).

Em virtude do direcionamento das informações que reforçam a fragmentação no debate público entre posições rivais, a involução democrática, a deterioração política e a necessidade compatibilizar liberdade de expressão com outros direitos fundamentais e com interesses e valores coletivos juridicamente relevantes, o Estado busca caminhos de regulação sobre a atividade dessas empresas de modo a atribuir responsabilidades e obrigações às gestoras dos fluxos comunicacionais de acordo com padrões de moderação de conteúdo conhecidos no mundo offline e com garantia de direitos processuais.

Nesse sentido, Evelyn Douek (2022) argumenta que a resolução de casos individuais tem sido priorizada como forma de governança de conteúdo do ambiente online, mas, segundo ela, os avanços na regulação dos discursos online dependem de uma compreensão sistêmica, de fluxo em rede, de escala, da necessidade de vigilância em massa e das possibilidades do tecnológicas.

Diante disso, a perspectiva a ser compreendida nesse trabalho é que a filtragem algorítmica não é um mecanismo de moderação de conteúdo pontual porque suas dimensões guardam correspondência com a velocidade e a escala do discurso online mas também não é um modelo infalível, no entanto, é capaz de contingenciar com rapidez e em tempo real a disseminação viral de fake news disparadas por milícias digitais, minimizando os efeitos danosos e frustrando a construção de cenários artificiais, diminuindo a circulação de conteúdo desinformativo e abrindo espaço no feed de notícias para informações verificadas e balanceadas e de fontes profissionais.

Apesar de sua eficiência, a filtragem algorítmica não é precisa e há estratégias para burlá-la como a substituição de algumas letras por números ou acentos gráficos dentro de uma palavra, permanecendo compreensível a leitura, mas escapando dos filtros algorítmicos que procuram excluir fake news e desinformação, por exemplo. Como são expressões matemáticas e buscam os exatos termos programados independentemente do sentido ou dos sinônimos, os bots cumprem tarefas automatizadas e, apesar de fazerem isso de maneira acurada, não são seres pensantes e podem ser driblados também por formas tecnológicas. Marcos Quiles esclarece como isso ocorre:

Entretanto, conhecendo os mecanismos usados pela plataforma, grupos maliciosos podem aplicar técnicas inversas para burlar os respectivos sistemas. Nesse contexto, diversas abordagens com os mais variados níveis de sofisticação têm sido empregadas. Por exemplo, para persuadir sistemas de moderação e de validação de

conteúdo, rotinas de camouflagem são utilizadas na transformação do texto, no qual letras ou sílabas específicas são substituídas dentro da mensagem, por exemplo, a letra “A” pelo número 4, como em V4CIN4. Embora simples, essa abordagem permite que textos indesejáveis numa sociedade democrática continuem sendo amplamente difundidos. (QUILES, 2022)

O autor ainda reforça que a desinformação online não se restringe a textos. Cada vez mais tem se aprimorado as deepfakes com imagens e montagens de rostos e sobreposição de vozes e isso tem tornado as fake news cada vez mais sofisticadas e perigosas porque houve incremento na dificuldade em demonstrar e convencer o usuário de que as fotos e os vídeos são falsos. Diante disso, é importante que o combate às fake news no ambiente online continue a se desenvolver tendo em vista que a atividade fraudulenta online tem avançado tecnologicamente e as formas analógicas de enfrentamento das fake news não têm obtido progresso no mesmo ritmo.

Apesar de não ser precisa e não atentar para as nuances, Evelyn Douek (2022) esclarece que existe uma compensação da filtragem algorítmica pela escala e velocidade desse tipo de moderação. O volume de dados online e a falta de tecnologia torna o Estado incapaz de moderar o conteúdo online (DOUEK, 2022, p. 06), no entanto, os critérios privados de filtragem algorítmica têm se mostrado opacos, inclusive porque as empresas gestoras dos fluxos comunicacionais se reservam a não esclarecer os métodos.

Nesse sentido, alegam que revelar a forma como identificam comportamentos coordenados em rede, como o das milícias, implica em munir os detratores de informações capazes de se desviarem dos padrões de moderação comportamental (DOUEK, 2022). Num vídeo (META, 2018), o chefe da política de segurança cibernética do Facebook tenta explicar como identificam comportamento coordenado inautêntico, mas é evasivo.

A perspectiva a ser compreendida é que “a internet não muda o que é a liberdade de expressão, entretanto altera as condições sociais nas quais as pessoas falam e isso traz à frente questões que se encontravam dormentes” (VALENTE, 2020, p.36). Por ser um ambiente dinâmico, fluido, cuja moderação de conteúdo lato sensu evolui de maneira acelerada para se compatibilizar com os avanços tecnológicos dos detratores, os controles online necessitam de parcerias público-privadas em situações inéditas em que o positivismo rígido de processos legislativos cede lugar a soluções tecnológicas mais rápidas e eficazes.

No entanto, a legalidade como medida de previsibilidade dos controles continua essencial no contexto democrático assim como transparência e imparcialidade do contingenciamento. As fake news assumem novas formas de exteriorização que não se

compatibilizam com argumentação evasiva nem excessivamente detalhista tendo em vista que uma caracterização pormenorizada e sem discricionariedade acaba por se tornar obsoleta inviabilizando adaptações ao ambiente online que é disruptivo e em acelerada transformação.

A revolução digital que transpôs as barreiras de tempo e de espaço, que forjou um novo homem que interage com máquinas tipo *smartphones* como extensão do próprio corpo, que provoca contínuo reaprendizado e disrupção tecnológica, reconfigurou conceitos, reprogramou diretrizes governamentais, acrescentou novos players na vida democrática e exigiu um modo diferente para fazer frente ao expansionismo das Big Tech, desnaturalizando a internet e as relações em rede. A comunicação online, a tendência de digitalização, a globalização e a ascensão econômica e transnacional das empresas de tecnologia que fazem a gestão dos fluxos e que transformaram em empresas de mídia social impõem uma nova forma de instrumentalizar o direito.

Diante disso, estuda-se uma forma eficiente de filtragem da desinformação online, mais especificadamente das fake news disseminadas por milícias digitais, sem necessariamente fazer avaliação caso a caso, partindo da análise de um controle sistêmico, explicitada as razões da decisão, o contingenciamento pode ser aplicado de forma automatizada sem a notificação particularizada de todos os envolvidos ou julgamento individualizado por um corpo imparcial.

4 INFLUÊNCIAS INTERNACIONAIS JURISPRUDENCIAIS E LEGISLATIVAS DA DINÂMICA VIRTUAL NO BRASIL E COMO AS FORMAS DE COMBATE ÀS FAKE NEWS TÊM EVOLUIDO NO ÂMBITO DO PODER JUDICIÁRIO, ESPECIALMENTE NO TSE

A importância do poder da fala, a capacidade e mesmo a necessidade de se exprimir socialmente é intrínseco a natureza humana, desde os tempos mais remotos. Na Antiguidade, Aristóteles (2017), filósofo grego, definiu o homem como animal político, sendo o dom do discurso o traço distintivo em relação a outros animais que vivem reunidos, ou seja, a fala seria a capacidade de transmitir, discernir e valorar situações e sentimentos e não meramente emitir sons. Timothy Ash (2017, p. 92), parafraseando Aristóteles, diz que “[...] o poder da fala é aquilo que nos distingue dos outros animais e, por enquanto, dos computadores”. Nesse sentido, da pré-história à modernidade, a liberdade de pensamento e de manifestação – liberdade de expressão - é referida como traço característico da condição humana, intrínseco, direito entranhado à sua existência, sendo desse modo uma expressão de sua dignidade.

No Brasil, a ADPF 130/DF (BRASIL, 2009), julgada em abril/2009 pelo STF, declarou a Lei de Imprensa como não recepcionada pela CF/88. A decisão foi importante marco na jurisprudência da liberdade de pensamento e de expressão, que foi qualificada, no voto do relator ministro Ayres Britto, como sobre-direito, tendo estabelecido uma preferência *prima facie* da liberdade de expressão em detrimento de outros direitos da personalidade, entendendo-a como norma-regra de aplicação imediata e que se impõe em toda e qualquer situação em concreto, “a salvo de qualquer restrição em exercício”, implicando um dever de não intervenção estatal, sujeitando-se apenas as disposições figuradas na própria Constituição.

O ministro relator Carlos Ayres Brito votou pela procedência da ADPF e expressou o entendimento da primazia *prima facie* das liberdades de pensamento e de expressão nos seguintes termos:

É hora de uma primeira conclusão deste voto e ele reside na proposição de que a Constituição brasileira se posiciona diante de bens jurídicos de personalidade para, de imediato crivar a primazia ou procedência: **a das liberdades de pensamento e de expressão lato sensu (que ainda barca todas as modalidades de criação e de acesso à informação, esta última em sua tríplice compostura, conforme reiteradamente explicitado)**. Liberdades que não podem arredar pé ou sofrer antecipado controle antecipado nem mesmo por força do Direito-lei, compreensivo este das próprias emendas à Constituição, frise-se. Mais ainda, liberdades reformadamente protegidas se exercitadas como atividade profissional ou habitualmente jornalística e como atuação de qualquer dos órgãos de comunicação social ou de imprensa. (BRASIL, 2009)

O julgamento da ADPF 130/DF pelo STF, em abril de 2009, no qual a Lei de Imprensa foi declarada não recepcionada pela CF/88, é um exemplo da diversidade de concepções sobre a liberdade de expressão. Apesar de todos os ministros entenderem a importância desse direito, discordam nas razões dos votos em vários aspectos sobre a proteção que deva ser conferida à liberdade expressão.

Sem pretensão de esgotar o julgado e sua complexidade, mas apenas de abordar notas importantes dos votos dos ministros, tem-se: o ministro Menezes Direito entendeu que dignidade da pessoa humana e liberdade de expressão são exigências de sobrevivência da sociedade democrática. Dignidade como precedência sobre a liberdade de expressão. Ellen Gracie discordou da primazia *prima face* da liberdade de expressão, arguida pelo relator. Ministro Gilmar Mendes não excluiu a possibilidade de limitações à liberdade expressão mediante lei frente a outros direitos constitucionais. Joaquim Barbosa, defendeu uma atuação positiva como fonte libertadora da liberdade de expressão. Para Celso de Mello, a liberdade de expressão não admite intervenção *a priori*, é fundamento da ordem democrática. Marco Aurélio foi o único que votou pela improcedência de pedido porque defendia que outra legislação substituísse a Lei de Imprensa para evitar que o pronunciamento judicial estabelecesse o vácuo legislativo. A ministra Carmem Lúcia arguiu que a liberdade expressão é garantia essencial à dignidade.

Vários documentos históricos e internacionais prestígiam a reconhecem a liberdade de expressão como direito essencial ao homem. Pode-se mencionar a Revolução Francesa, em 1789, cujo lema foi “liberdade, igualdade e fraternidade”. A cláusula geral de proteção em face do poder público que foi reconhecida no movimento iluminista europeu também foi positivada na Primeira Emenda da Constituição Americana, bem como na Declaração Universal de Direitos do Homem, aprovada pela assembleia da ONU em 1948, e na Convenção Europeia dos Direitos do Homem, art. 10, que passou a vigorar em 1953.

Assim, a liberdade de expressão, apesar de não ser direito novo, não tem seus exatos termos definidos porque está em contínua construção. Isso ocorre porque a extensão dos direitos do homem é processo histórico, caracterizado por conquistas contra poderes estabelecidos (BOBBIO, 2020). Originalmente, a liberdade de expressão reconheceu como ameaça a figura hierárquica superior, derivada de uma relação de poder verticalizada em que as críticas ao governo e aos autoridades públicas poderiam ensejar retaliações e perseguições. Desse modo,

a liberdade de manifestar insatisfações continua a ser um exercício legítimo, essencial a democracia e ao pluralismo de ideais.

Os Estados Unidos da América (EUA) foram o país percussor a decidir sobre os contornos da liberdade de expressão e, como não havia e nem há entendimento uniforme sobre a extensão da liberdade de expressão, os limites sobre o exercício legítimo desse direito foram construídos ao longo dos anos, pela jurisprudência da Suprema Corte. Os casos foram moldando uma interpretação da Primeira Emenda abrangente, conduzindo a jurisprudência que privilegia a liberdade de expressão em detrimento de outros direitos como dignidade, igualdade e privacidade. Para Robert A. Sedler (2016) isso ocorreu porque as garantias do Bill of Rights foram redigidas em termos gerais e abrangentes, dependentes de interpretação pelos ministros da Suprema Corte, num momento histórico que as normas de direito internacional eram ainda incipientes e não haviam sido incorporadas ao direito constitucional americano.

Em virtude do exposto, é oportuno trazer à tona as primeiras concepções jurisprudenciais sobre a temática inserindo os julgamentos na linha do tempo, a forma como foi amadurecida a compreensão da liberdade de expressão e as razões pelas quais o modelo americano tem uma vertente mais protetiva da autonomia individual e dá mais ênfase ao conteúdo jurídico subjetivo da liberdade de expressão.

Vitor Amaral Medrado (2019) fez uma interessante digressão sobre as restrições à liberdade de expressão nos EUA. Segundo ele, no período colonial, além do sistema de licenciamento prévio que autorizava ou não as publicações, vigorava a lei de Difamação Sediciosa que criminalizava a publicação de material desrespeitoso à Igreja, ao Estado ou aos seus representantes e nem mesmo à comprovação de veracidade afastava a incidência legal. Segundo o autor, esse instituto foi reeditado sob o nome de lei de Sedição, em 1798. Assim, mesmo após a independência, foi aprovada regulação que tornava a manifestação contra os poderes estabelecidos um crime federal, no intuito de frear ideais da Revolução Francesa nos EUA.

Ainda de acordo com o autor, os debates políticos sobre a dita lei e a derrota dos federalistas, tendo Thomas Jefferson assumido a Presidência e anistiado todos os que tinham sido condenados com base na lei de Sedição, abriu espaço para a posição libertária da liberdade de expressão. Em 1919 ocorreram os primeiros julgamentos na Suprema Corte que trouxeram as luzes iniciais para a concepção americana. Apesar de ter sido voto vencido, o ministro Holmes Jr oxigenou as compreensões do tribunal ao introduzir novas abordagens.

No caso *Schenck vs United States*, ele trouxe à lume a fórmula do “perigo claro e presente” como a condição para a punição à liberdade de expressão, utilizando o exemplo do alarme de incêndio no teatro lotado como situação justificável de restrição à liberdade de expressão, nesse sentido nem mesmo a posição mais libertária em favor da liberdade expressão protegeria quem falsamente gritasse fogo num teatro lotado, levando ao pânico e caos. No mesmo ano, no caso *Abrams vs United States*, o justice refinou a ideia original para “perigo claro e iminente” reforçando a necessidade do potencial danoso da manifestação para incidir penalização. A interpretação de um trecho de seu voto popularizou a expressão livre mercado de ideias. Em 1925 no caso *United States vs Schwimmer*, novamente em voto vencido, Holmes Jr reconhece o direito à liberdade de expressão ainda que seja para de difundir ideias heterodóxicas.

Outros julgamentos influenciaram uma posição da corte mais favorável a liberdade de pensar e manifestar. Vitor Amaral Medrado (2019) relaciona casos nos anos seguintes que envolvem a temática da liberdade de expressão e liberdade de imprensa e que ampliaram sua extensão como o de *Stromberg vs California* de 1931 e o de *Near vs Minnessota* de 1936 em que os ministros vencidos em 1927 no caso *Grosjean vs American PressCo* mudaram de opinião.

O emblemático *New York vs Sullivan* em 1964 mais uma vez reconheceu a liberdade de expressão como fundamento da democracia. Segundo a decisão não há necessidade de provar a veracidade dos fatos para impedir a responsabilização. No caso, a liberdade de imprensa prevaleceu sobre os direitos do servidor que se sentiu difamado, mesmo não tendo sido comprovados os fatos reportados pelo jornal. No mesmo julgamento foi declarada inconstitucional a lei de Sedição, desenvolvendo a linha de uma atuação negativa do Estado.

Além desses, o caso *Time, Inc vs Hill* de 1967 contribuiu para discussões sobre o direito a privacidade, preterido no caso em detrimento da liberdade de expressão. O julgamento de *Branzburg vs Haves* de 1972 reconheceu aos jornalistas o direito de não testemunhar em juízo sob pena dessa exigência comprometer a liberdade de expressão e da imprensa garantida pela Primeira Emenda.

Apesar da concepção libertária da liberdade de expressão ter se fixado como preponderante no entendimento da Suprema Corte Americana, Gustavo Binenbojm e Caio Mário da Silva Pereira Neto (2005) fazem referência a teoria democrática da liberdade de expressão na interpretação da Primeira Emenda Americana, em contraposição a posição libertária. Segundo os autores essa concepção está mais centrada na figura do receptor do discurso, de modo a priorizar a formação do cidadão para que possa livremente ser informado

sobre assuntos de interesse geral e forme sua convicção. É uma corrente doutrinária diferente da prevalente nos EUA porque ao invés de privilegiar o direito individual pondera a liberdade de expressão sob a dimensão transindividual e está mais centrada no debate público robusto cujo propósito é a liberdade política.

A posição que se consolidou nos EUA é amplamente favorável à liberdade de expressão, mas ainda assim esse direito não foi compreendido como absoluto, admitindo restrições como o perigo real e iminente. Os críticos mais contundentes a respeito das limitações à liberdade de expressão, argumentam que esse direito deve ser o mais abrangente possível para promover o debate vigoroso de ideias, essencial ao sistema político republicano. Nesse sentido, justificam suas posições nas lições de Jonh Stuart Mill (2018), ocorre que mesmo para o autor o princípio do dano autoriza contenções a liberdade individual de se autogovernar. Desse modo, quando a conduta produz dano a outra pessoa se justifica o poder de regular e de impedir alguém de dizer ou fazer algo.

No campo do debate, esse filósofo defende que todas as ideias deveriam ser admitidas, mesmo as menos factíveis, e justifica esse posicionamento na circunstância de que, progressivamente, o debate público construiria a verdade e o consenso. Segundo o autor (2018, p. 45) “a verdade de uma opinião é parte de sua utilidade”, dessa forma, o objetivo do debate não seria apenas desvendar a melhor escolha pelo confronto de argumentos e experiências divergentes, mas também o aprendizado coletivo pela reconciliação e combinação de opostos de tal modo que todas as ideias divergentes deveriam ser encorajadas até que se atinja o ajuste entre as opiniões (MILL, 2018, p. 86).

Assim, a verdade como um processo em constante construção está sempre submetida aos argumentos e a revisões e, por isso, é instigada a se defender numa batalha hostil. No caso das fake news e das dimensões inéditas de tempo e espaço do ambiente online, a batalha hostil enfrentada pela verdade nunca se finda porque há camadas e mais camadas de fake news que se sobrepõem ao fato, onde discussões infundáveis não convergem para um ponto comum, uma posição negociada entre opiniões diferentes. As fake news tumultuam o processo de tomada de posição especialmente quando disseminadas no ambiente online porque permanecem acessíveis e aptas a produzir os efeitos da desinformação tendo em vista não esmaecerem nem serem esquecidas, estando sempre vívidas na memória virtual.

As colocações de Mary Anne Frank (2011) a respeito da ideias de Stuart Mill trazem uma nova perspectiva sobre o autor. Segundo ela, a teoria de Mill não parte do pressuposto de que todas as manifestações são válidas, inclusive porque o Mill estabelece como limite à

liberdade o dano ao outro. O argumento de Mill, de que mesmo as opiniões falsas podem contribuir para a formação da verdade pelo confronto de visões diferentes, é uma postura inclusiva do autor, de abertura para minorias, para que excluídos sejam ouvidos, no sentido de que suas manifestações e participações na esfera pública favorecem ao pluralismo e a oxigenação dos debates democráticos.

Assim, segundo Mary Anne Frank (2011), Mill percebeu a necessidade proteger minorias contra a tirania social, ou seja, o autor captou a tendência de que posições majoritárias intimidem e silenciem vozes discordantes, preocupando-se muito mais com comportamentos sociais opressores do que com a ação governamental de contingenciamento. A disseminação de fake news por milícias digitais pode ser comparável a tirania social temida pelo autor, onde um pequeno grupo de perpetradores com o uso instrumentos tecnológicos promove ataques simultâneos para impor uma narrativa ou construir artificialmente um cenário que pareça majoritário suplantando vozes divergentes e silenciando minorias.

Por isso, Mary Anne Frank (2011) defende que Mill não se posicionou contrariamente à regulação do discurso, especialmente quando essa interferência é feita para oportunizar mais igualdade na participação social. Ao considerar manifestações divergentes como tendo igual valor e, portanto, dignas da mesma consideração no embate de ideias, Mill dá oportunidade para minorias e excluídos se manifestarem. Portanto, para Mary Anne Frank (2011) nem mesmo a teoria de Mill sobre liberdade de expressão daria suporte ao discurso de assédio ou que se distancie dos propósitos cívicos de um diálogo aberto.

Por óbvio, a regulação do discurso online segundo princípios e fundamentos democráticos acarretará restrições à liberdade de expressão dos violadores de direitos e disseminadores de fake news que intencionalmente agem para confundir e distorcer a opinião pública ao difundir fatos sabidamente falsos, construindo falsos cenários no intuito de obter ganho ou causar prejuízo.

A preponderância da liberdade de expressão face a outros direitos individuais não é uma trajetória retilínea e ao longo da história teve avanços e retrocessos até que a posição liberal se estabelecesse como majoritária nos EUA. Os Estados Unidos são o país sede da maioria das empresas de tecnologia e isso explica o motivo pela qual as empresas gestoras dos fluxos comunicacionais online absorveram e se ajustaram às concepções americanas, inclusive com incentivo legal da Seção 230 do ADC, a Lei de Decências das Comunicações, que contém cláusula de imunidade aos provedores por lhes reconhecer apenas como intermediários de

conteúdo. Além disso, como são considerados meros intermediários, são proibidos de proceder edições às manifestações de terceiros em seus domínios.

No entanto, em virtude das práticas de modulação e moderação de conteúdo online a percepção americana sobre as gestoras dos fluxos comunicacionais como meras empresas de tecnologia, intermediadoras, fica prejudicada. Ao dispor sobre a seleção, organização e estruturação da informação essas empresas induzem a significados e interpretações específicos, influenciando ou mesmo determinando opiniões e comportamentos. Nesse sentido Manuel Castells (2020, p. 494), refere-se a fluxo não apenas como elemento de organização social, mas como “expressão dos processos que *dominam* nossa vida econômica, política e simbólica.”, portanto, aquele que tem a gestão de fluxos acaba por dominar algum aspecto da vida, razão pela qual operar a gestão do fluxos comunicacionais representa poder das empresas sobre os usuários.

A concepção americana de conduzir o ciberespaço segundo a lógica de mercado que se rege por regras internas de *laissez faire* foi um estímulo ao desenvolvimento dessas empresas e de suas tecnologias em expansão, no entanto, acabou por permitir que essas empresas crescessem sob um sistema frouxo de regulação e se eximindo de responsabilidades diante do trânsito de informações. Nesse sentido, Castells (2020, p. 70) é preciso ao destacar o papel do Estado, inclusive como regulador, diante das inovações tecnológicas, nos seguintes termos:

O que de ver guardado para o entendimento da relação entre a tecnologia e a sociedade é o papel do Estado, seja interrompendo, seja promovendo, seja liderando a inovação tecnológica, é um fator decisivo no processo geral, à medida que expressa e organiza as forças sociais dominantes em um espaço e uma época determinada. (CASTELLS, 2020, p. 70)

Do mesmo modo, Habermas (2018) se refere aos modelos de política americana que se contrapõem: a liberal e a republicana. Segundo ele, na visão liberal, o cidadão é definido pelos direitos subjetivos que detém frente ao Estado e aos demais cidadãos e, assim, os direitos, inclusive a liberdade de expressão, são assimilados como direitos de defesa em face do Estado, que permite ao cidadão perseguir seus interesses dentro de limites estabelecidos pela lei, protegidos de intervenções do poder público. Nesse sentido, o processo político visa controlar o poder estatal, de modo que o agregado de interesses privados forma a vontade popular e submete o Estado às deliberações públicas. Na política liberal, os direitos subjetivos dos cidadãos são direitos negativos diante do Estado e por isso se sobrepõem ao Estado e estipulam os limites do poder estatal na consecução dos fins coletivos. Segundo a visão liberal, a economia, as leis de mercado e interesses privados é que norteiam a conduta estatal e administração pública.

Ainda segundo Habermas (2018, p. 398), a visão republicana de política, que se contrapõe à liberal, elege a solidariedade como fonte de integração social, não mais o interesse próprio ou a dinâmica do mercado. Nesse contexto, os cidadãos se guiam por razões éticas, reconhecendo entre eles uma relação de dependência mútua, sendo uma “associação de parceiros do direito livres e iguais”. Desse modo, a noção de disputa de direitos entre os cidadãos cede ao compartilhamento de direitos em igualdade de condições e o processo político republicano passa a ser orientado ao entendimento em que os direitos são compreendidos como liberdades positivas de participação e comunicação a serem garantidas pelo Estado. Assim, na visão republicana, a vontade geral é orientada pelo poder comunicativo em que os cidadãos se entendem sobre o que corresponde ao interesse comum.

Assim, o modelo político define a posição estatal frente aos direitos. Ao transpor as discussões sobre a liberdade de expressão para o contexto da comunicação online, percebe-se que o argumento liberal da “competição do mercado” de combater a desinformação com mais informação no sentido de propagar a verdade para descortinar as fake news pode ter o efeito inverso do desejado. Primeiro porque as fake news não se esgotam depois porque a competição não é equilibrada, os fluxos comunicacionais priorizam fake news e conteúdo desinformativo, desestabilizando qualquer disputa que pretenda racional porque o homem passa a competir com máquinas e algorítmicos que tem uma capacidade de processamento muito acima da cognição humana tanto para produzir fake news quanto para sofisticar as técnicas de fabricação a exemplo das deep fakes, sendo ainda mais difícil desvendar a falsidade.

Em contraposição a corrente americana que prestigia a liberdade de expressão como direito individual quase que absoluto, a Europa e mais especificadamente a Alemanha conduzem uma leitura de vanguarda, mais atenta à repercussão social e coletiva e aos danos imensuráveis da manifestação online haja vista a longevidade e o alcance das postagens online que perpetuam os efeitos nocivos e ampliam a visibilidade a cada instante. Nesse sentido, a construção doutrinária e jurisprudencial da Alemanha abriu espaço para a promulgação da NetzDG, uma lei que criminaliza determinadas condutas online impondo às empresas gestoras dos fluxos comunicacionais obrigações de excluir conteúdos logo que tenham conhecimento das situações tipificadas na lei, além do dever de estabelecer um canal para reclamações e denúncias, acessível e online, com o dever de processá-las e dar resposta em prazo razoável estabelecido na lei.

Assim, a Alemanha rompe com a ideia da autorregulação das empresas e a substitui pela autorregulação regulada em que a lei, e não o mercado, é que norteia a atividade das

empresas que fazem a gestão dos fluxos comunicacionais online, impondo obrigações e responsabilidades, estabelecendo diretrizes no ciberespaço segundo uma posição compartilhada na produção das normas. No Brasil, o projeto de Lei 2630/20 que ficou conhecido como lei das fake news e atualmente se propõe a ser a Lei Brasileira de Liberdade, Responsabilidade e Transparência segue a mesma orientação da NetzDG, estipulando obrigações para as empresas gestoras dos fluxos comunicacionais online (BRASIL, 2020a).

As regras de mercado segundo um modelo liberal prestigiavam a autorregulação dessas empresas de modo a expandirem negócios e a gerirem os fluxos comunicacionais online mediante moderação e modulação de conteúdo com finalidades lucrativas, se apropriando das manifestações, lucrando com compartilhamentos, coleta, venda e processamento de dados, direcionando publicidade e fazendo prognósticos preditivos sobre comportamentos dos usuários, priorizando o modelo de negócios em detrimento a direitos da personalidade e de bens coletivos juridicamente relevantes como a democracia que se viam a mercê de ataques e atividades fraudulentas, criando embaraços a compreensão de realidade compartilhada, manipulando a opinião pública e influenciando decisões democráticas.

Desse modo, a autorregulação não se adequa ao Estado de Direito porque submete milhões de pessoas a padrões de comportamento estabelecidos unilateralmente por contratos individuais e implementados de maneira opaca, segundo finalidades lucrativas, suplantando o interesse público, a representação política plural, os ritos e procedimentos democráticos previamente estabelecidos para tomada de decisão que afetam populações inteiras.

A Alemanha ascende também na construção do princípio da autodeterminação informativa a partir do art. 2º § 1º da sua lei fundamental. Num estudo sobre o tema, Laura Mendes (2020) traz uma sequência de julgados e demonstra como foi desenvolvido o entendimento a respeito desse dispositivo invocando como direitos autônomos a liberdade de ação e o direito da personalidade, apesar de terem a mesma raiz normativa. A liberdade de ação como sendo a garantia de livre comportamento do indivíduo. E o direito a personalidade como sendo o direito que a pessoa tem de decidir o que será ou não tornado público a seu respeito.

A seguir são relacionados os principais julgados trazidos pela autora e que conduziram a percepção da autonomia informativa e da teoria horizontal dos direitos fundamentais contra terceiros. Nas decisões nos casos do microcenso (informações sobre lazer e viagens de férias) e do divórcio (autos de divórcio remetido para investigador de processo administrativo disciplinar) o Tribunal Constitucional Alemão discute a esfera de proteção privada do indivíduo perante terceiros e a necessidade de consentimento para ter acesso a informações de cunho

íntimo, no entanto, decide que as informações sobre viagem não têm caráter sigiloso e por isso não trouxeram prejuízo a personalidade. Sobre a remessa dos autos do divórcio considerou inconstitucional, por ter avaliado as informações conjugais como íntimas e de acesso restrito ao público.

Ainda segundo a autora, em 1973, num outro julgado, discutiu-se sobre o direito geral de personalidade e a necessidade de proteção da personalidade diante das ameaças tecnológicas. A situação versava sobre o uso de uma gravação obtida secretamente num processo de sonegação fiscal e houve a ponderação entre o interesse público e a esfera privada. O Tribunal Constitucional Federal entendeu que os direitos fundamentais são garantias de defesa e não havia, no caso, interesse público relevante que justificasse a violação a esfera privada.

O outro julgamento importante mencionado no estudo referido tratou sobre a proteção da personalidade em virtude do documentário sobre assassinato do soldado “Lebach”. No caso, um dos criminosos argumentava que a exibição poderia afetar sua ressocialização. O Tribunal trata não apenas da proteção da esfera privada do criminoso, mas do reconhecimento do direito geral de personalidade, de modo abstrato, diante da sociedade. Em 1980, no julgamento conhecido como “Eppler”, o Tribunal procedeu uma explicação minuciosa do direito de personalidade como sendo o direito que a pessoa tem de decidir o que será ou não tornado público a seu respeito (ideia de autodeterminação informativa), concepção aprimorada pela evolução jurisprudencial e amparada no Direito Constitucional.

A construção do direito geral de personalidade dá-se no intuito de proteger o indivíduo diante das inovações da sociedade moderna. O foco da fundamentação não é mais a esfera privada, tendo em vista as dificuldades em identificar qual conteúdo está inserido no âmbito privado de cada um, já que esse conteúdo varia de pessoa para pessoa. A questão agora é trazida para apreciação judicial, do que pode ou não ser tornado público, sendo o conceito de dados pessoais aberto e flexível.

Dando seguimento a seleção jurisprudencial alemã feita por Laura Mendes (2020) no sentido de contextualizar o direito a autodeterminação informativa, ela menciona o emblemático recenseamento de 1983. No caso, se discute a proteção de dados e os riscos da personalidade diante das novas modalidades de processamento algorítmico, de tal modo que não é mais a natureza privada ou íntima da informação que determina sua proteção porque qualquer dado pessoal passa a ser relevante a depender da finalidade da análise. Desse modo, a proteção dos dados pessoais passa a ser abrangente de forma a abarcar coleta, armazenamento, utilização e transmissão.

O princípio da autodeterminação informativa, segundo a autora (MENDES, 2020), deriva da interpretação conjugada do direito geral de personalidade (art. 2º, § 1º, LF) e da dignidade da pessoa humana (art. 1º, § 1º da LF), sendo, desse modo, uma continuidade da proteção abstrata da personalidade em que o indivíduo passa a ter o controle sobre seus dados pessoais e pode decidir quais informações fornece, para quem e com qual objetivo.

De acordo com o estudo referido, o Tribunal Constitucional Alemão entendeu no caso do recenseamento de 1983, devem ser tomados cuidados para anonimização dos dados pessoais além de observância aos deveres de sigilo de modo a evitar que o indivíduo se torne um objeto de informação perante o Estado. Apesar do direito a autodeterminação informativa ser entendido como uma garantia constitucional, o Tribunal Constitucional Alemão entende que não é ilimitado e não haveria impedimento a realização do censo desde que tomadas as cautelas referidas porque as informações colhidas são relevantes para a formação de um retrato social.

No tocante a aplicação do direito à autodeterminação informativa nas relações privadas, Laura Mendes (2020) relaciona 02 casos, sendo o primeiro em 1991 em que o locatário do imóvel omitiu a informação de que era interditado e alegou seu interesse no sigilo a interdição. Nesse caso, o Tribunal Constitucional Alemão, em sede de recurso, considerou a dimensão objetiva do direito geral à personalidade que estaria em risco diante de rotulação social e da impossibilidade de alugar uma moradia. O segundo caso referido por ela (MENDES, 2020) foi em relação as causas da incapacidade em contratos de seguros, em que o segurado deveria conceder uma autorização genérica conferindo poderes para que a seguradora buscasse informações sobre a causa da incapacidade profissional com dispensa geral da confidencialidade. A decisão do Tribunal Constitucional foi no sentido de reconhecer a dimensão objetiva dos direitos fundamentais como uma autoproteção informativa do indivíduo diante de outro ente particular.

Segundo Laura Mendes (2020) outro fator que pesou nas decisões foram as disparidades das partes contratantes de modo a perturbar o princípio da paridade contratual sendo o consentimento meramente fictício a darem causa a atuação do Estado no sentido de proteger o direito à autodeterminação informativa. A decisão do Tribunal Constitucional Alemão no caso do seguro é de vanguarda porque se fundamenta não apenas em relação a autonomia privada do direito obrigacionista mas de aplicação do direito a personalidade cuja interpretação evoluiu para a proteção ao tratamento de dados pessoais.

A jurisprudência alemã influenciou o Brasil na elaboração do art 2º, II, da Lei Geral de Proteção de Dados, Lei nº 13.709 de 2018 (BRASIL, 2018a). Nesse sentido, os usuários passam

a ter maior autonomia em relação a seus dados, estabelecendo o dever de comunicação de como estão sendo utilizados, dando-lhes a oportunidade de revogar os consentimentos, restringindo a negociação dos dados como mercadoria, podendo servir como fundamento para pedidos de exclusão de informações falsas.

Diante da revolução tecnológica, a maioria das pessoas se informa por notícias online e a qualidade das informações nos buscadores são importantes ferramentas para enfrentamento das fake news, por isso informações relevantes e verificadas sobre candidatos e candidaturas logo nos primeiros resultados de busca diminuem a desinformação e a facilidade das verificações tem potencial para interromper a corrente de compartilhamentos.

Ao tomar conhecimento de forma fácil e rápida sobre a falsidade de determinado conteúdo, o usuário pode romper a transmissão da desinformação e passar a ser um ponto de esclarecimento para outros usuários. Desse modo, a autodeterminação informativa pode dar ao candidato o direito de atualização das informações sobre sua trajetória política bem como a prioridade para conteúdo verificado, de modo a otimizar as informações nos primeiros resultados de busca ao invés de priorizar as fake news. Nesse sentido, há uma inversão na ordem de apresentação das notícias e informações no intuito de preservar direitos de terceiros no ambiente online.

Os dados pessoais rotineiramente capturados e negociados pelas empresas gestoras dos fluxos comunicacionais e essenciais para a formação de bancos de dados sobre os usuários receberam a proteção constitucional no Brasil em 2022, pela Emenda Constitucional 115, que incorporou o inciso LXXIX ao art. 5º da CF/88. Com esse dispositivo, a proteção de dados pessoais no meio online passa a ser um direito e garantia individual fundamental, dando ao indivíduo maior controle sobre a extração, apropriação e tratamento online de seus dados pessoais, reforçando os direitos a segurança e a privacidade.

No Brasil, as práticas online impactaram mudanças não apenas quanto a proteção de dados, mas também quanto aos contornos da liberdade de expressão. O combate as fake news disseminadas por milícias digitais foi especialmente determinante numa nova compreensão sobre a liberdade de expressão no país. O entendimento de primazia prima face da liberdade de expressão estabelecido no julgamento da ADPF 130/DF em 2009 pelo STF sofre uma mudança diante do cenário tecnológico e das repercussões inéditas sobre os efeitos das fake news online.

As eleições americanas de 2016 e a revelação escândalo da Cambrigde Analytics em 2018 foram outro marco histórico no enfrentamento da desinformação e evidenciaram como a

democracia e o resultado das eleições pode ser suscetível as inovações digitais, seja pelo uso dos dados pessoais que direciona conteúdo e anúncios com mensagens subliminares ou pela influência das fake news disparadas em massa que atordoam os cidadãos semeando dúvida quanto a credibilidade dos candidatos, a falsidade de suas propostas ou mesmo desconfiança nas autoridades e processos eleitorais.

Diante das ameaças aos processos eleitorais e às regras do jogo democrático, em 2017 o Tribunal Superior Eleitoral – TSE, órgão de cúpula da Justiça Eleitoral no Brasil formado por 07 ministros, dentre eles 03 ministros do STF, 02 ministros do STJ e 02 advogados, criou o Conselho Consultivo sobre a Internet para combater as fake news através da Portaria TSE nº 949/2017. O foco eram as fake news e a preocupação de que repetisse nas eleições brasileiras de 2018 o que tinha sido observado no cenário internacional (BRASIL, 2017a).

As discussões abordavam sobre a necessidade de criação de cartilhas e campanhas de conscientização para a população, elaboração de manuais de procedimentos para os juízes eleitorais, criação de um ambiente virtual (site ou aplicativo) para recebimento de sugestões de atuação e denúncias sobre fake news e análise das propostas de resolução do TSE a fim de apresentar colaborações (BRASIL, 2022a). No mesmo ano também foi debatido o uso de robôs para propagação de conteúdo de baixa qualidade bem como a necessidade de educação midiática dos cidadãos no intuito de diminuir a circulação de fake News (BRASIL, 2023b).

Como medidas de combate à desinformação online, em junho de 2018 foram assinados termos de cooperação com 03 entidades representativas do setor de comunicação - Associação Brasileira de Emissoras de Rádio e Televisão (Abert), a Associação Nacional de Jornais (ANJ) e a Associação Nacional de Editores de Revistas (ANER) – para desestimular e coibir a produção e a disseminação de conteúdos falsos ou enganosos durante o período eleitoral. Também foram firmados acordos com representantes do Google e Facebook no sentido de comprometimento com a verdade dos fatos e de não proliferação de fake news, diminuindo a possibilidade de replicação de práticas aptas a distorcer a liberdade de voto e as escolhas dos cidadãos (BRASIL, 2022b).

Ocorre que nas eleições de 2018, além das fake news na propaganda eleitoral sobre partidos, candidatos e propostas, houve nas palavras da ministra Rosa Weber a partir do mês de agosto, “um movimento paralelo de direcionamento maciço de ataques à Justiça Eleitoral com a divulgação em larga escala de notícias falsas visando ao descrédito da instituição e de seus integrantes bem como a suspeição do sistema eletrônico brasileiro de votação”, trecho do

discurso da ministra proferido em 2019 na abertura do Seminário Internacional Fake News e Eleições (SEMINÁRIO INTERNACIONAL..., 2020).

A concepção até então era a de menor interferência na liberdade de expressão, tanto que em 21 outubro 2018, o ministro Luís Felipe Salomão nega liminar para retirada de conteúdos publicados em redes sociais contra Fernando Haddad, com fundamento no art. 33 da Resolução 23.551/2017 do TSE (BRASIL, 2017b), já revogada, em que o caput tinha a seguinte redação: “Art. 33. A atuação da Justiça Eleitoral em relação a conteúdos divulgados na internet deve ser realizada com a menor interferência possível no debate democrático (Lei nº 9.504/1997, art. 57-J).”

Segundo notícia reportada no site do TSE (BRASIL, 2017b), na decisão, o ministro acrescentou que, em redes sociais, “é possível, por meio da utilização do espaço destinado aos comentários da publicação, o questionamento e a contraposição pelos usuários da internet ou pelo próprio ofendido acerca da veracidade do conteúdo postado, alertando aos demais para a falsidade das informações”. Nesse sentido, mostra-se visível a preponderância do debate de ideias em detrimento da exclusão de conteúdo, deixando a cargo dos próprios usuários o descortinamento da verdade e a elucidação dos fatos.

Os ataques à credibilidade do sistema eletrônico de votação, à segurança da urna eletrônica e do resultados das eleições em 2018 foram construindo uma nova compreensão da liberdade de expressão diante dos efeitos expandidos no ciberespaço com dimensões inéditas de tempo, espaço e memória virtual. As repercussões e danos do exercício arbitrário da liberdade de expressão mostraram que as medidas de controle constitucionalmente expressas como direito de resposta e reparação pecuniária eram inadequadas e insuficientes para frear os efeitos das fake news porque, enquanto online, continuavam a gerar prejuízos. Então o contingenciamento passou a ser mecanismo adotado não somente pelas próprias empresas para preservar seus domínios diante das expectativas dos usuários e dos anunciantes, mas também como medida judicial para excluir conteúdo violador de direitos, restringindo a liberdade de expressão dos perpetradores.

Assim, em 2019, entre as eleições presidenciais de 2018 e 2022, no âmbito do Poder Judiciário Brasileiro foi instaurado o inquérito policial nº 4781, conhecido como inquérito das fake news, que tramita em sigilo, tendo como relator o ministro Alexandre de Moraes, designado pelo presidente do STF para condução do feito, onde foram determinadas, por medida liminar, ordens de exclusão de conteúdo e cancelamento de perfis, buscas e apreensões.

O objeto da investigação é a apuração de “fatos relacionados a fake news e denúncias caluniosas, ameaças e infrações revestidas animus *calumniandi*, *diffamandi* e *injuriandi* que atingem a honorabilidade e a segurança do Supremo Tribunal Federal, de seus membros e familiares” (p. 06/380) e do “vazamento de informações e documentos sigilosos, com o intuito de atribuir e/ou insinuar a prática de atos ilícitos por membros da Suprema Corte, por parte daqueles que tem o dever legal de preservar o sigilo; e a verificação da existência de esquemas de financiamento e divulgação em massa nas redes sociais, com o intuito de lesar ou expor a perigo de lesão a independência do Poder Judiciário e ao Estado de Direito”, conforme consta na ADPF 572/DF (BRASIL, 2020b, p. 29/380).

O Partido Rede Sustentabilidade ajuizou Ação de Descumprimento de Preceito Fundamental, ADPF 572/DF, em face da Portaria GP nº 69 de março de 2019 do STF que determinou a abertura do inquérito nº 4781 (conhecido como inquérito das fake news) alegando lesão ou ameaça de lesão a preceito fundamental da liberdade pessoal em face a garantia do devido processo legal (art. 5º, LIV, CF/88), da dignidade da pessoa humana (art. 1º, III, CF/88), da prevalência dos direitos humanos (art. 4º, II, CF/88), da legalidade (art. 5º, II CF/88) e da vedação a juízos e tribunais de exceção (art. 5º, XXXVII, CF/88). No acórdão, em junho de 2020, o plenário do STF julgou improcedente a ação e declarou constitucional a Portaria 69/2019 e o artigo 43 do RISTF:

(...) por maioria, conheceu da arguição de descumprimento de preceito fundamental, converteu o julgamento da medida liminar em julgamento definitivo de mérito e, nos limites do processo, diante do incitamento do fechamento do STF, de ameaça de morte ou prisão de seus membros, de apregoada desobediência a ordens judiciais, julgou totalmente improcedente o pedido nos termos expressos em que foi formulado ao final da petição inicial para declarar a constitucionalidade da Portaria 69/2019 enquanto constitucional o artigo 43 do RISTF, nas específicas e próprias circunstâncias de fato com esse ato exclusivamente envolvidos, nos termos do voto do Relator e dos votos proferidos, vencido o ministro Marco Aurélio. (BRASIL, 2020b, p. 29/380).

Apesar do ministro-relator da ADPF 572/DF, Edson Facchin, arguir em seu voto que defende a liberdade *prima facie*, ressalva que, em casos concretos, a manifestação desse direito pode se tornar abusiva.

Mesmo com a preponderância que a liberdade de expressão assume em nosso sistema de direitos, e de sua “posição de preferência” [preferred position], seu uso em casos concretos pode se tornar abusivo. Neste sentido, podem-se agregar ao exercício legítimo da liberdade de expressão alguns condicionantes que balizem a aferição de responsabilidades civis e penais.” (BRASIL, 2020b, p. 40/380)

Nas razões do voto, o ministro faz referência ao Pacto Internacional dos Direitos Civis e Políticos em que o art. 19 impõe condições para que se admitam restrições ao exercício da liberdade de expressão pelo Estado que deverão estar fixadas em lei e desde que necessárias

para assegurar a proteção aos direitos ou à reputação de outrem, a proteção da segurança nacional, a ordem pública ou à saúde e a moral públicas, nos seguintes termos:

A evolução dos variados sistemas de proteção dos direitos humanos, ao lado das tendências dominantes de práticas estatais sugerem que a restrição à liberdade de expressão deve ser permeada por alguns subprincípios. Em seu relatório do ano de 2017, o Relator Especial do Conselho de Direitos Humanos das Nações Unidas para a Promoção e a Proteção do Direito de Liberdade de Opinião e Expressão afirmou:

Diferentemente da proibição incondicional de ingerência na opinião, o artigo 19(3) do Pacto [Internacional dos Direitos Civis e Políticos] impõe três condições segundo as quais o exercício do direito à liberdade de expressão pode estar sujeito a restrições por parte dos Estados. Essas condições devem ser levadas a efeito de maneira estrita (cf. a Observação Geral nº 34 do Comitê de Direitos Humanos, parágrafos 21 a 36). O art. 19(3) dispõe que o exercício do direito à liberdade de expressão envolve deveres e responsabilidades especiais e pode estar sujeito a certas restrições, que deverão, contudo, estar expressamente fixadas em lei e ser necessárias para: a) assegurar a proteção aos direitos ou à reputação de outrem; a proteção da segurança nacional, a ordem pública ou à saúde e moral públicas. O art. 20 do Pacto também estabelece que toda propaganda em favor da guerra e toda apologia ao ódio nacional, racial ou religioso que constitua incitação à discriminação, à hostilidade ou à violência estarão proibidos por lei (**Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression**. Organização das Nações Unidas: A/71/373, 2016). (BRASIL, 2020b, p. 29/380)

Do texto transcrito acima, reproduzido no voto do ministro-relator, Edson Facchin, não há margem para interpretação judicial sobre quais seriam as hipóteses admitidas de restrições da liberdade de expressão, já que está expressamente consignada a obrigatoriedade de estarem escritas em lei (ou seja, estabelecidas pelo Legislativo, segundo rito próprio dos debates parlamentares)..

Ocorre que no Brasil, não existe uma lei que defina o que pode ou não ser dito, e a razão é que ao particular é permitido tudo o que não seja proibido, princípio da legalidade, art. 5º, II, da CF/88, sendo mesmo inadmissível se pretender uma lei que regule todas as situações de manifestação da liberdade de expressão. No entanto, é possível uma lei para definir limites, ou seja, positive hipóteses em que cabível restrição à liberdade de expressão. Assim, como não existe no direito brasileiro lei específica que traça as balizas, o ministro Edson Facchin, por interpretação conclui: “O desenho constitucional também inscreveu na República brasileira a proteção do Estado de Direito democrático e por isso mesmo dos Poderes instituídos, razão pela qual a preservação das instituições é essencial na democracia representativa.” (BRASIL, 2020b, p. 51/380)

Por sua vez, o ministro Alexandre de Moraes ao proferir o voto no julgamento da ADPF 542/DF, arguiu a defesa do Poder Judiciário expressamente em 3 situações, ao menos, repete: “Não há democracia sem Poder Judiciário” (BRASIL, 2020b, p. 87/380; 125/380; 136/380). O transcurso do tempo e dos posicionamentos da corte é que esclarecerão se foi

estabelecida um nova primazia e que ensejaria novas limitações à liberdade de expressão. Assim, a liberdade de expressão não estaria garantida quando utilizada contra aos membros do Poder Judiciário, merecendo o Poder Judiciário uma defesa prima facie em relação à democracia? No contexto do inquérito das fake news, que tramita em segredo de justiça, foram relacionadas postagens contra o STF e seus ministro e familiares, algumas realmente ameaçadoras e com caráter incendiário, outras, no entanto, não tem, a princípio, a natureza criminosa, mas de protesto, mas como tramita em segredo de justiça, o quê foi considerado de fato ofensivo e marginal ainda não foi tornado público.

Em seu voto na ADPF 572/DF, o ministro Alexandre de Moraes faz referência ao julgamento de outra ação, ADI 4.451/DF (BRASIL, 2018b), em que foi relator e faz a transcrição de afirmações feitas naquele voto em 2018 para justificar seu posicionamento, no entanto, acrescenta uma restrição que não constava no julgamento referido e que altera a compreensão anterior, incerta na parte final, último parágrafo inclusive com grifo do próprio ministro.

A instauração de procedimento investigatório não se direcionou, assim como o próprio inquérito não se direciona a ampla liberdade de expressão, consagrada constitucionalmente e inúmeras vezes reafirmada por essa SUPREMA CORTE.

Conforme afirmei na ADI 4451, de minha relatoria:

- 1.A Democracia não existirá e a livre participação política não florescerá onde a liberdade de expressão for ceifada, pois esta constitui condição essencial ao pluralismo de ideias, que por sua vez é um valor estruturante para o salutar funcionamento do sistema democrático.
2. A livre discussão, a ampla participação política e o princípio democrático estão interligados com a liberdade de expressão tendo por objeto não somente a proteção de pensamentos e ideias, mas também opiniões, crenças, realização de juízo de valor e críticas a agentes públicos, no sentido de garantir a real participação dos cidadãos na vida coletiva.
- 3.São inconstitucionais os dispositivos legais que tenham a nítida finalidade de controlar ou mesmo aniquilar a força do pensamento crítico, indispensável ao regime democrático. Impossibilidade de restrição, subordinação ou forçosa adequação programática da liberdade de expressão a mandamentos normativos cerceadores durante o período eleitoral.
- 4.Tanto a liberdade de expressão quanto a participação política em uma Democracia representativa somente se fortalecem em um ambiente de total visibilidade e possibilidade de exposição crítica das mais variadas opiniões sobre os governantes.

Porém, a Constituição Federal consagra o binômio LIBERDADE com RESPONSABILIDADE; não permitindo de maneira irresponsável a efetivação de abuso no exercício de um direito constitucionalmente consagrado; não permitindo a utilização da “liberdade de expressão” como escudo protetivo para a prática de discursos de ódio, antidemocráticos, ameaças, agressões, infrações penais e toda a sorte de atividade ilícitas. LIBERDADE DE EXPRESSÃO NÃO É LIBERDADE DE AGRESSÃO LIBERDADE DE EXPRESSÃO NÃO É LIBERDADE DE DESTRUIÇÃO DA DEMOCRACIA, DAS INSTITUIÇÕES E DA HONRA ALHEIA. (BRASIL, 2018b)

O parágrafo final transcrito e grifado no voto do ministro Alexandre de Moraes na ADPF 572/DF, não consta do pronunciamento manifestado na ADI 4.451/DF, sendo acrescentada como nova restrição, acrescentada como fundamento das razões do voto da ADPF 572/DF.

A ADPF 572/DF julgou a constitucionalidade do conhecido inquérito das fake News, (ainda sigiloso). Dentre outros comportamentos, se apuram aqueles manifestados online, motivo pela qual foram determinadas medidas restritivas inclusive de bloqueio de perfil de usuário de redes sociais, com a justificativa de que as disseminações de mensagens investigadas partiam das referidas contas (BRASIL, 2020c). Em seu voto no ADPF 572/DF, a ministra Carmem Lúcia, registra:

E demonstrou bem o Ministro Edson Fachin, e enfatizou muito, em seu voto, o Ministro Alexandre de Moraes, os atos, objeto da investigação questionada, não constituíram em nenhum momento, pelo menos do que aparenta até aqui e que se comprovou no que já foi a providência adotada, objeto nem de liberdade de expressão, nem de, principalmente, liberdade de imprensa. O que se teve aqui foi um conjunto de atos que precisavam de ter a sua investigação quanto aos autores e às suas finalidades, que são exatamente contrários aos princípios democráticos que a Constituição estabelece.

A ministra também reitera o argumento presente em quase todos os votos, a defesa do Judiciário, e nesses termos se refere:

Nesse sentido portanto, quando, como autodefesa democrática, pela utilização de instrumentos legítimos, dos princípios e das regras que são adotadas, um dos quais é o da separação e harmonia entre os Poderes, se formula a proposta ou, neste caso, a imposição de uma ferramenta, de um instrumento que garante com que um dos Poderes, neste caso, foi o Poder Judiciário, não possa ser, de alguma forma, comprometido pela ação de uns poucos, por ações, portanto, antijurídicas, inconstitucionais e antidemocráticas.

O Poder Judiciário foi a via eleita para dirimir controvérsias sobre o exercício abusivo ou não da manifestação da liberdade de expressão online (art. 19 da Lei 12.965/2014 – Marco Civil da Internet – adiante referida pela sigla MCI). Desse modo, casuisticamente tem se traçado os contornos do exercício ilícito online. Ocorre que as decisões judiciais, na valoração de princípios constitucionais, através da ponderação de interesses em jogo, trazem incertezas sobre os limites gerais da liberdade de expressão.

No caso específico da ADPF 572/DF, os votos dos ministros Alexandre de Moraes e Carmem Lúcia parecem estabelecer uma primazia de proteção do Poder Judiciário como garantia para a democracia. Assim, a liberdade de expressão e a própria democracia poderia ser preteridas diante de manifestações contra o Poder Judiciário, seus membros e familiares, estabelecendo uma diferença no tratamento da liberdade de expressão entre os demais cidadãos.

Em continuidade as medidas contra a desinformação, no âmbito da Justiça Eleitoral, nas eleições para prefeitos, governadores e parlamentares em 2020, o TSE criou a página fato ou fake “para esclarecer sobre informações relacionadas ao processo eleitoral” com “divulgação de notícias checadas, recomendações e conteúdos educativos” (BRASIL, 2020d, [n. p.]). Segundo a mesma notícia do site do TSE, o Programa de Enfrentamento à Desinformação conta com mais de 70 instituições entre partidos políticos e entidades públicas e privadas para combate à desinformação.

Nas eleições presidenciais do Brasil de 2022, além da página “Fato ou Fake” que conta com parceiros para checagem de informações, foram produzidos vídeos no YouTube, no canal do TSE, sobre o sistema eletrônico de votação. Além disso, o TSE disponibilizou um assistente virtual, acessível por smartphone, denominado “Tira Dúvidas do TSE no WhatsApp”, um chatbot que simula uma conversa humana e esclarece dúvidas em tempo real. Esse canal proporciona uma interação dos usuários com o TSE, “aumentando o número de cadastrados para receber checagem sobre notícias falsas, bem como oferecer serviços da Justiça Eleitoral e aprimorar a navegabilidade para os usuários.” (BRASIL, 2022c)

Foi criada também uma ferramenta para denúncias no “Sistema de Alerta de Desinformação” em que os usuários podem apontar números de telefones suspeitos de disparo de mensagens em massa. Além disso, as eleições foram acompanhadas por comissão internacional de observação eleitoral, na qual foi elaborado relatório pela Universidade de Complutense de Madrid. De acordo com notícia no site do TSE, as ações do tribunal de combate à desinformação foram elogiadas, no entanto, foi sugerido:

uma maior integração entre as áreas jurisdicional e administrativa, embora tenha sido realizado um trabalho conjunto, e continuação do processo de fortalecimento da segurança jurídica, evitando alterações de última hora nas normas e consolidando as mudanças por meio de regulamentações, não apenas jurisdicionais (BRASIL, 2022d, [n. p.]).

Apesar de todo esse preparo para enfrentar as fake news no período eleitoral, segundo o Presidente do TSE, houve um aumento de denúncias que motivaram a edição da Resolução 23.714/2022 entre o 1º e o 2º turnos das eleições de 2022. A fundamentação para a proposta e aprovação da Resolução 23.714/2022 do TSE foi o incremento da desinformação em relação ao processo eleitoral, a rapidez que as fake news circulam e os efeitos de difícil de reparação especialmente diante da proximidade do pleito eleitoral que colocam em risco a estabilidade do panorama político e o exercício de direitos fundamentais como o acesso a informações verdadeiras.

Abaixo transcrito o trecho do relatório em que o Presidente do TSE justifica, em seu voto, a necessidade de novo instrumento jurídico, Resolução 23.714/2022 do TSE, apresentando a escalada repentina dos níveis de desinformação:

O Termo de Informação encaminhado pela Assessoria Especial de Enfrentamento à Desinformação (AEED/TSE) noticia, em linhas gerais, forte incremento no volume de conteúdos falsos em torno do processo eleitoral, conforme os seguintes dados: a) crescimento de 1.671% (mil seiscentos e setenta e um por cento) no volume de denúncias de desinformação encaminhadas às plataformas digitais, em comparação com as eleições de 2020, sendo que cerca de 47% (quarenta e sete por cento) dos apontamentos surgiram no breve período de doze dias após o final de semana das eleições; b) necessidade de publicação mais de 130 (cento e trinta) novas matérias com desmentidos e esclarecimentos sobre casos de desinformação grave que ganharam ampla visibilidade após o primeiro turno de votação; c) recebimento de uma média diária de 506 denúncias de desinformação contra as eleições, o que corresponde a um aumento de 992% na taxa de apontamentos normalmente tratada no âmbito da Assessoria Especial de Enfrentamento à Desinformação. Além disso, ficou constatado o aumento dos episódios de violência política, que alcança 436% (quatrocentos e trinta e seis por cento) comparado a 2018, ano da última eleição presidencial, conforme estudos publicados pelas ONGs Justiça Global e Terra de Direitos (Disponível em: [Violência Política e Eleitoral no Brasil - Terra de Direitos]. Acesso: 17.10.2022.), que possivelmente surgem como desdobramento natural da polarização exacerbada ocasionada pelo fenômeno da desinformação. (BRASIL, 2022e, p. 03)

A Resolução 23.714/2022, publicada entre o 1º e 2º das eleições presidenciais de 2022, aumentou as multas para o caso de descumprimento de ordem judicial, concedeu maior poder de polícia ao Presidente do TSE para combate a desinformação relacionada ao processo eleitoral, diminuiu prazos e possibilitou a exclusão de perfis e contas, além de suspensão de serviços das plataformas, a seguir:

RESOLUÇÃO Nº 23.714, DE 20 DE OUTUBRO DE 2022.

Dispõe sobre o enfrentamento à desinformação que atinja a integridade do processo eleitoral.

O TRIBUNAL SUPERIOR ELEITORAL, no uso de suas atribuições legais e regimentais,
RESOLVE:

Art. 1º Esta Resolução dispõe sobre o enfrentamento à desinformação atentatória à integridade do processo eleitoral.

Art. 2º É vedada, nos termos do Código Eleitoral, a divulgação ou compartilhamento de fatos sabidamente inverídicos ou gravemente descontextualizados que atinjam a integridade do processo eleitoral, inclusive os processos de votação, apuração e totalização de votos.

§ 1º Verificada a hipótese prevista no caput, o Tribunal Superior Eleitoral, em decisão fundamentada, determinará às plataformas a imediata remoção da URL, URI ou URN, sob pena de multa de R\$ 100.000,00 (cem mil reais) a R\$ 150.000,00 (cem e cinquenta mil reais) por hora de descumprimento, a contar do término da segunda hora após o recebimento da notificação.

§ 2º Entre a antevéspera e os três dias seguintes à realização do pleito, a multa do § 1º incidirá a partir do término da primeira hora após o recebimento da notificação.

Art. 3º A Presidência do Tribunal Superior Eleitoral poderá determinar a extensão de decisão colegiada proferida pelo Plenário do Tribunal sobre desinformação, para outras situações com idênticos conteúdos, sob pena de aplicação da multa prevista no

art. 2º, inclusive nos casos de sucessivas replicações pelo provedor de conteúdo ou de aplicações.

§ 1º Na hipótese do caput, a Presidência do Tribunal Superior Eleitoral apontará, em despacho, as URLs, URIs ou URNs com idêntico conteúdo que deverão ser removidos.

§ 2º A multa imposta em decisão complementar, proferida na forma deste artigo, não substitui a multa aplicada na decisão original.

Art. 4º A produção sistemática de desinformação, caracterizada pela publicação contumaz de informações falsas ou descontextualizadas sobre o processo eleitoral, autoriza a determinação de suspensão temporária de perfis, contas ou canais mantidos em mídias sociais, observados, quanto aos requisitos, prazos e consequências, o disposto no art. 2º.

Parágrafo único. A determinação a que se refere o caput compreenderá a suspensão de registro de novos perfis, contas ou canais pelos responsáveis ou sob seu controle, bem assim a utilização de perfis, contas ou canais contingenciais previamente registrados, sob pena de configuração do crime previsto no art. 347 da Lei nº 4.737, de 15 de julho de 1965 - Código Eleitoral.

Art. 5º Havendo descumprimento reiterado de determinações baseadas nesta Resolução, o Presidente do Tribunal Superior Eleitoral poderá determinar a suspensão do acesso aos serviços da plataforma implicada, em número de horas proporcional à gravidade da infração, observado o limite máximo de vinte e quatro horas.

Parágrafo único. Na hipótese do caput, a cada descumprimento subsequente será duplicado o período de suspensão.

Art. 6º É vedada, desde quarenta e oito horas antes até vinte e quatro horas depois da eleição, a veiculação paga, inclusive por monetização, direta ou indireta, de propaganda eleitoral na Internet, em sítio eleitoral, em blog, em sítio interativo ou social, ou em outros meios eletrônicos de comunicação da candidata ou do candidato, ou no sítio do partido, federação ou coligação (art. 7º da Lei n. 12.034, de 29 de setembro de 2009).

§ 1º Verificado descumprimento da vedação a que se refere o caput, o Tribunal Superior Eleitoral, em decisão fundamentada, determinará às plataformas a imediata remoção da URL, URI ou URN, sob pena de multa de R\$ 100.000,00 (cem mil reais) a R\$ 150.000,00 (cem e cinquenta mil reais) por hora de descumprimento, a contar do término da primeira hora após o recebimento da notificação.

§ 2º O descumprimento do disposto no caput configura realização de gasto ilícito de recursos eleitorais, apto a determinar a desaprovação das contas pertinentes, sem prejuízo da apuração do crime previsto no art. 39, § 5º, inciso IV, da Lei n. 9.504, de 30 de setembro de 1997.

Art. 7º O disposto nesta Resolução não exclui a apuração da responsabilidade penal, do abuso de poder e do uso indevido dos meios de comunicação.

Art. 8º Fica revogado o art. 9º-A da Resolução TSE nº 23.610, de 2019.

Art. 9º Esta Resolução entra em vigor na data de sua publicação.

Brasília, 20 de outubro de 2022.

Referida Resolução estabelece a obrigatoriedade das plataformas excluïrem conteúdos especificados em prazos de 02 horas sob pena de multas, esse prazo é ainda diminuído, apenas 01 hora após a notificação, quando a determinação judicial ocorre na entre a antevéspera e os três dias seguintes à realização do pleito. Diferentemente da NetzDG que impõe multas em virtude do descumprimento reiterado sobre exclusões de conteúdo ou negativa repetida de apresentação dos relatórios.

O artigo 3º da Resolução 23.714/2022, concede ao Presidente do TSE o poder para estender os efeitos de uma decisão do colegiado, a conteúdos desinformativos idênticos replicados em outros endereços eletrônicos. A extensão dos efeitos de decisão colegiada do

Tribunal pelo Presidente do TSE de modo a excluir conteúdo desinformativo que tenha identidade ao já decidido pela corte em julgamento anterior, a par de ser justificada na segurança jurídica, coerência, efetividade e agilidade das decisões judiciais que combatem a desordem informativa, implica em mitigação das garantias processuais das pessoas frente ao poder público como a do devido processo legal, como contraditório e ampla defesa.

A Resolução 23.714/2022 acaba por instituir um controle sistêmico online, a ser implementado de maneira automatizada, tendo em vista o prazo fulminante de horas para cumprimento da decisão pelas plataformas notificadas sob pena de vultosas multas. Além disso, decisões monocráticas sem a sessão pública de um julgamento do colegiado, que habitualmente é acompanhado pelas partes e por advogados, acabam por dificultar a compreensão dos contornos da liberdade de expressão que são construídos pelo Poder Judiciário.

Nas situações de remoções de conteúdo desinformativo sem comunicação prévia, a extensão dos efeitos da decisão do Tribunal mediante a determinação do Presidente do TSE de exclusão de conteúdo idêntico em outros endereços eletrônicos se equipara à moderação de conteúdo proativa em que as empresas gestoras dos fluxos comunicacionais online fazem o controle dentro de seus domínios segundo normas de segurança, civilidade e comportamento, de acordo com os termos de serviço ou diretrizes da comunidade.

Ocorre que os usuários, pessoas físicas ou jurídicas, que tiveram conteúdo excluído diante de ordem emanada do Poder Judiciário têm garantias constitucionalmente asseguradas perante o Estado sobre atos oriundos do Executivo, Legislativo e Judiciário. No entanto, na circunstância de extensão dos efeitos de decisão do colegiado por ordem do Presidente do TSE as garantias do devido processo legal não são observadas porque as exclusões não são comunicadas previamente nem oportunizada defesa a quem teve a liberdade de expressão contingenciada por extensão dos efeitos da decisão.

Diante disso, se ao Poder Judiciário, que tem obrigações processuais constitucionalmente previstas na CF/88, para evitar o exercício abusivo de suas funções, mitiga as garantias processuais sob o argumento da efetividade, coerência e celeridade de suas decisões, os limites e responsabilidades que se pretende impor às empresas gestoras dos fluxos comunicacionais, pessoas jurídicas privadas dentro de seus espaços particulares e submetidas a regra geral de legalidade, no sentido de evitar o silenciamento das minorias e diminuição do pluralismo na mídia se fragilizam, deixando os usuários cada vez mais suscetíveis a arbitrariedades.

A tendência da moderação de conteúdo online operada pelas empresas é assegurar um mínimo de garantias processuais aos usuários, ainda que de forma diferente das prescrições constitucionais que se aplicam em relação ao poder estatal, no entanto, a margem de discricionariedade ampliada, trazida na Resolução 23.714/2022 do TSE, abre caminho para que as moderações de conteúdo sejam cada vez mais decisões com ampla margem para contingenciamento da liberdade de expressão.

Caracterizar identidade de conteúdo é difícil e porque há variações na escrita, com diferentes formas de redigir ainda que sejam sobre um mesmo conteúdo, de modo que um jogo de palavras diferentes traz significados diferentes sobre um mesmo fato e nem sempre podem ser classificados como fake news. Desse modo, a extensão dos efeitos da decisão pelo Presidente do TSE trata-se do poder conferido para fazer avaliações monocráticas sobre um tema decidido pelo tribunal, em que a identidade de conteúdo referida na Resolução será, na prática, uma similitude trazendo o risco da margem de discricionariedade desaguar em subjetivismos e arbítrios especialmente tendo em vista os prazos fulminantes contados em poucas horas sob pena de multas vultosas, prejudicando até mesmo a recorribilidade e reversibilidade da decisão de contingenciamento.

A moderação de conteúdo online mediante a exclusão de fake news disparadas por milícias digitais de modo automatizado mediante filtragens algorítmicas com o uso de inteligência artificial tem menor margem de erro do que o cumprimento de determinação do Presidente do TSE porque o controle algorítmico pelas empresas se opera sobre a mesma mensagem, com idêntica escrita segundo um mesmo padrão de disparos. A partir de um controle interno, as empresas procedem exclusões de uma mesma mensagem, com a mesma redação, disseminadas por perfis em atividade coordenada, feitas dentro do ambiente de cada empresa, não sendo uma ordem geral que pode incidir sobre toda a internet, a depender da quantidade de plataformas notificadas.

A remoção de conteúdo para estender os efeitos de um julgamento colegiado do Tribunal segundo determinação do Presidente do TSE concentra poderes que vão muito além do poder de polícia, ao atribuir o contingenciamento na figura do Presidente do TSE, a liberdade de expressão pode ficar sujeita ao entendimento pessoal da autoridade que centralizar a interpretação e o reconhecimento da identidade de conteúdo exigida na Resolução. A atividade de julgar, interpretar e reconhecer a identidade de conteúdo, apesar de ser pretensamente imparcial, não é neutra.

O artigo 4º da Resolução 23.714/2022 faz referência a produção sistemática de desinformação, entendidas como notícias falsas ou informações descontextualizadas, sobre o processo eleitoral, autorizando a suspensão temporária de perfis, contas ou canais mantidos em mídias sociais. Essa é uma medida muito mais grave, em termos de contingenciamento da liberdade de expressão, do que a exclusão de conteúdo. É de se observar que a situação se aplica, nos termos da redação, quando o contexto da desinformação se refere ao processo eleitoral, ou seja, assuntos relacionados à falsas suspeitas sobre autoridades eleitorais como juízes e ministros, ao sistema eletrônico de votação, à segurança da urna eletrônica e à confiabilidade do resultado das apurações dentre outros. Pelo que se pode depreender, não se adequa à situação de fake news que lançam desconfiança sobre candidaturas ou candidatos.

O artigo 5º da Resolução 23.714/2022 aumenta o poder de polícia do Presidente do TSE ao conferir-lhe a possibilidade de suspensão dos serviços de plataformas, impactando a vida de milhões de usuários no país. A suspensão dos serviços das plataformas já tinha sido objeto de decisão do ministro Edson Facchin na ADPF 403/SE (BRASIL, 2021c), em 28 de maio de 2020, cujo julgamento pelo STF ainda está pendente, consulta realizada em 12 de março de 2023.

Na ADPF 403/SE, o ministro-relator Edson Facchin, entendeu que a sanção de suspender serviço do WhatsApp, prevista no art 12. III, do MCI, era uma penalidade a ser aplicada somente pela Autoridade Nacional e não pelo Poder Judiciário. Desse modo, a aplicação da penalidade do art. 12, III, da Lei 12.965/2014, pelo juiz foi considerada inconstitucional, sem redução de texto. Segundo a decisão, nos casos em que o juiz determinou a interceptação telemática de conversas criptografadas do WhatsApp, não poderia determinar a suspensão do serviço diante do descumprimento da decisão judicial porque, de acordo com o voto, essa sanção cabe à Autoridade Nacional e não ao Judiciário. A ADPF 403/SE teve o julgamento suspenso em virtude de pedido de vista do Ministro Alexandre de Moraes, em 28 de maio de 2020. De acordo com consulta processual no site do STF realizada em 13 de março de 2023, os autos foram devolvidos em 06 de março de 2023, mas não houve colocação em pauta para continuidade de julgamento.

Ao tratar sobre o viés institucional que pesa sobre a liberdade de expressão, Cass Sunstein (2018) argumenta que o “teste de perigo claro e iminente” é uma excelente resposta em caso de risco institucional generalizado, porque as avaliações sobre custos e benefícios costumam ser distorcidas em virtude de interesses próprios em jogo. Apesar de não ser a melhor solução para todas as situações, Cass Sunstein (2018) argumenta que, pelo risco de dano à

liberdade de expressão, o “teste de perigo claro e iminente” tem se apresentado com a melhor dentre outras possíveis.

Segundo ele, na avaliação de custos e benefícios, muitas vezes os funcionários públicos tendem, de maneira consciente ou não, a se isolar das críticas, preocupando-se mais em proteger-se do que à sociedade dos danos, confundindo o interesse em manter sua autoridade e proteger seu poder em detrimento da estabilidade, paz e violência.

In authoritarian nations, the point is clear. Even in the United States, history speaks volumes.¹³ Public officials have often argued, and often apparently believed, that restrictions on speech were necessary to avoid serious harm, certainly in connection with war and national security threats but also outside of those contexts. In retrospect, their arguments look implausible; it seems clear that they were seeking to censor or punish ideas and perspectives with which they disagreed. At the time, however, their claims were highly credible. In the heat of the moment, it is all too easy to convince the public and even oneself that suppression of speech is necessary to avoid harm, even if that is entirely false. In recent years, it is easy to think of examples, whether they come from Russia, China, Cuba, or Turkey. On this view, citizens in authoritarian nations Free Speech 201 desperately need the protection of the clear and present danger test, or something close to it—and citizens in democracies do as well. (SUNSTEIN, 2018, p. 200/201)

Nesse sentido, a avaliação das redes é de perdas e ganhos, custos e benefícios com a moderação de conteúdo, pesando expectativa dos usuários e fake news online. No âmbito judicial, o contingenciamento de conteúdo é avaliado segundo um critério de perigo real e iminente para evitar que interesses particulares se sobreponham ao interesse público. Apesar da avaliação do perigo real e iminente poder ser fantasiosa e imprecisa é menos suscetível a sucumbir a interesses privados do que o critério de custos e benefícios.

Assim, nas situações em que as empresas gestoras dos fluxos comunicacionais fazem os próprios controles, avaliam perdas e ganhos balanceando os lucros entre excluir ou manter uma fake news. Diante disso, a precisão das filtragens de fake news, segundo expectativas de usuários e anunciantes que não querem estar associados a ambientes tóxicos e depreciativos, ficam prejudicadas tendo em vista que fake news também geram engajamento e lucros. Apesar disso, o critério de controle de fake news disseminadas por milícias digitais não deve ser desprezado, ao contrário, deve ser estimulado e regulado para que seja cada vez mais eficiente.

Num contexto social dividido, polarizado, o combate as fake news deve ser enfrentado sem escolha ideológica, evitando diferenças no tratamento entre as forças que disputam no jogo democrático, direita ou esquerda, republicanos ou democratas. Norberto Bobbio (2019, p. 103) faz advertência sobre a diferença entre as exclusões de elementos subversivos à organização política e a exclusão de todas as formas de organização política menos a oficial, ou seja, “a que impõe não apenas as regras do jogo, mas também o único modo através do qual se deve jogar”.

O presente trabalho não diferencia o combate das fake news segundo atores ou a corrente ideológica que a mensagem adere, defende a exclusão pela comprovação de falsidade do conteúdo, independentemente do teor ideológico, associando-se outro critério de moderação que é padrão de comportamento articulado em rede. A imparcialidade, como tratamento uniforme e igual, é medida que se impõe no controle das fake news sob o risco de ferir mortalmente a democracia, ainda que tenha por escopo defendê-la.

De acordo, Steven Levitsky e Daniel Ziblatt (2018, p. 81) argumentam que as iniciativas governamentais para trazer fôlego à democracia acabam por sufocá-la:

A erosão da democracia acontece de maneira gradativa, muitas vezes em pequeníssimos passos. Tomado individualmente, cada passo parece insignificante – nenhum deles aparenta de fato ameaçar a democracia. Com efeito, as iniciativas governamentais para subverter a democracia costuma ter um verniz de legalidade. Elas são aprovadas pelo Parlamento ou julgadas constitucionais por supremas cortes. Muitas são adotadas sob o pretexto de diligencie algum objetivo público legítimo - e mesmo elogiável - , como combater a corrupção, ‘limpar’ as eleições, aperfeiçoar a qualidade da democracia ou aumentar a segurança nacional.

Dito isso, a insatisfação, as denúncias, o dissenso, a divergência política e mesmo a desinformação têm espaço no ambiente democrático porque a democracia é um regime de tensões latentes que se revezam no poder, então, os protestos, as reivindicações populares, o compartilhamento desavisado sem os cuidados médios de checagem e os erros de julgamento quanto aos fatos fazem parte do contexto. O estudo não pretende pôr fim às fake news ou a movimentos de protesto causando a supressão da liberdade de expressão como meio para prevenir ações populares porque o dissenso e o direito de associação para fins lícitos fazem parte da democracia, que é um regime plural.

Nesse sentido Bobbio (2019, p. 104) conclui: “Creio assim ter explicado por que vinculei o problema do pluralismo ao dissenso. Bem consideradas as coisas, apenas numa sociedade pluralista o dissenso é possível: mais ainda, não apenas é possível, mas necessário.” Assim, a crítica, a insatisfação, o protesto são meios legítimos numa democracia, ainda que desencadeados em rede, como o quê aconteceu na Primavera Árabe e nas manifestações estudantis no Brasil em junho de 2013. Do mesmo modo ocorrem com as fake news, fazem parte do contexto democrático. O estudo não pretende um ambiente online livre de falsidades ou mensagens descontextualizadas, apenas que no meio virtual haja uma reprodução das fake news na proporção em que se apresenta na realidade física.

O disparo de fake news por comportamentos milicianos coordenados em rede com o uso de robôs aciona os mecanismos de impulsionamento algorítmico e acaba por desequilibrar as forças que disputam o poder, corrompendo as regras do jogo democrático. Mensagens

disparadas por robôs e perfis falsos não são manifestação cívica da liberdade de expressão, mas uso disfarçado do poder econômico para impor um pensamento ou resultado através de mecanismos tecnológicos e artifícios subliminares, que acabam por dominar de forma inconsciente o pensar e o agir, manipulando cidadãos. A atividade miliciana facilita a constatação do ânimo doloso dos disseminadores e da natureza fraudulenta das fake news que visam a objetivos específicos.

No tocante, ao combate a desinformação no Brasil tem sido observada uma distinção no enfrentamento da desinformação, que tem se dividido em três frentes: fake news lançadas contra cidadãos, candidatos e partidos; ataques contra autoridades e instituições do Poder Judiciário, como o processo eleitoral; e mensagens odiosas como manifestação de violência. A desinformação que atinge a credibilidade no processo eleitoral, como a confiança no sistema de votação, a lisura das autoridades da Justiça Eleitoral, a segurança das urnas eletrônicas e a aceitação do resultado das eleições, tem merecido uma atuação mais enérgica, com medidas não aplicáveis para o controle da violência política ou das fake news contra candidatos, candidaturas ou cidadãos, nos termos da Resolução 23.714/2022, estabelecendo uma diferença no tratamento entre autoridades eleitorais e cidadãos, ainda que candidatos.

O perigo das medidas excetivas é se tornarem regra, as situações para extensão dos efeitos de decisão do TSE não decorrem de lei, mas de Resolução do TSE, que amplia o poder de polícia do Presidente do próprio tribunal. Além disso, a circunstância a ser avaliada de identidade de conteúdo com a decisão colegiada do TSE tem uma ampla margem de discricionariedade.

A forma de controle trazida pela Resolução 23.714/2022 se aproxima da teoria da democracia defensiva concebida na Alemanha por Karl Loewestein, cujas ideias já tinham sido abordadas na decisão do relator Edson Facchin nos autos da ADPF 572/DF em 2020, quando tratou sobre o inquérito das fake news, sendo um indicativo de que a tese do constitucionalista alemão já circulava entre ministros do STF desde 2020. No entanto, como é um modelo de moderação recém surgido no Brasil ainda é cedo para classificá-lo, mas a semente de um controle sistêmico para contingenciamento célere de desinformação, com ampla margem de discricionariedade, mediante uma autoridade central do Poder Judiciário foi inaugurada no país.

A avaliação, interpretação e julgamento da identidade de conteúdo desinformativo, ainda que com redação diferente está concentrada numa única autoridade, sem que o mesmo instrumento normativo tenha trazido a possibilidade de impugnação dessa decisão monocrática pelo plenário, em tempo igual, inviabilizando o direito de defesa e o contraditório de quem teve

o conteúdo excluído pela alegada identidade de conteúdo. Como resultado, o mais provável será a censura colateral tendo em vista que, para não ter os serviços interrompidos ou responder por multas vultosas, a empresa preferirá excluir conteúdo do que mantê-lo acessível.

O capítulo correlacionou a liberdade de expressão desde a concepção americana como direito negativo do cidadão diante do Estado até regulação público-privada inaugurada na Alemanha de autorregulação regulação dos fluxos comunicacionais online, em que a NetzDG o projeto de lei 2630/2020. No Brasil, as medidas de controle avançaram de uma não-intervenção estatal para formas mais ativas como acordos de cooperação com entidades privadas, estabelecendo a prioridade para informações verificadas no feed de notícias e no resultado nos buscadores até a concretização das exclusões de conteúdo por ordens judiciais, inclusive com a possibilidade de extensão dos efeitos dos julgamentos colegiados, que determinaram a remoção de conteúdo desinformativo, por decisões monocráticas, além de sanções como suspensão temporária dos perfis, contas, perfis multas, além da suspensão dos serviços das empresas gestoras dos fluxos comunicacionais online por desobediência.

5 ANÁLISE CONSTITUCIONAL DA FILTRAGEM ALGORÍTMICA DE FAKE NEWS DISSEMINADAS POR MILÍCIAS DIGITAIS

A comunicação online se opera de forma diferente da comunicação presencial, tem uma dinâmica própria controlada por empresas privadas segundo um modelo de negócios que subverte os valores da liberdade de expressão em razão dos incentivos econômicos criarem estímulos a proliferação de conteúdo que prende mais atenção do público em detrimento do caráter informativo ou de participação democrática. O ambiente de discussão online se rege por formas de comunicação midiaticizadas por algoritmos em que sociabilidade é codificada pela tecnologia e expedientes da realidade física não se encaixam às práticas virtuais.

Desse modo, a racionalidade tecnológica e social do meio digital em que as formas de comunicação estão baseadas na conectividade com combinações algorítmicas tendenciosas ao lucro e a interação social em regramentos informativos próprios e autônomos tem colocado em xeque a compreensão tradicional dos direitos fundamentais (WIELSCH, 2020). Por isso Wolfgang Hoffmann-Riem (2021) entende que “A princípio, é útil que as disposições sobre direitos fundamentais, mesmo que tenham uma longa tradição, sejam interpretadas dinamicamente ao longo do tempo, com o objetivo de garantir que suas premissas permaneçam relevantes diante de realidades em transformação.”

O estudo e a contextualização das filtragens algorítmicas de fake news disseminadas por milícias digitais perpassa pela necessidade de compreender os valores da liberdade de expressão essenciais ao desenvolvimento da personalidade do indivíduo - que se individualiza à medida que se expressa - e à manutenção do diálogo público pluralista em que as pessoas divergem, mas são permeáveis às trocas argumentativas. Nesse sentido é importante considerar a lógica do modelo de negócios que estrutura a comunicação online segundo finalidades lucrativas e que privilegia fake news para enfim analisar o controle das fake news através de mecanismos tecnológicos de moderação de conteúdo online.

No cenário de ampla participação online e controle incipiente, as fake news se multiplicaram e passaram a produzir efeitos negativos afetando direitos individuais e bens de interesse coletivo juridicamente relevantes como a democracia. Assim, apesar do ambiente de discussão online ser operado num espaço privado, a forma como o fluxo comunicacional circula traz consequências desastrosas para a democracia e ao exercício da cidadania (CALLEJÓN, 2022a).

Como ponto de partida para a análise sobre a exclusão de fake news disseminadas por mídias digitais, os valores da liberdade de expressão de Balkin (2020) dão a dissertação o referencial teórico para o exercício da liberdade de expressão no contexto democrático. Segundo o autor, os valores da liberdade de expressão servem à democracia política, à democracia cultural e ajudam a promover o crescimento e a disseminação do conhecimento.

Desse modo, Balkin (2020) explica que a democracia política é baseada em participação, receptividade e público informado. Nesse condão, o público indaga e o Estado responde sobre questões de interesse público de modo a que a informação prestada serve à formação da opinião pública. Segundo o autor (BALKIN, 2018), a governança do Facebook, Twitter e YouTube Facebook protege apenas direitos e as liberdades civis básicas e responde à opinião pública de maneira limitada.

Ainda de acordo com o autor (BALKIN, 2020), na democracia cultural, a liberdade de expressão permite a criação de significados que formam os indivíduos, dando oportunidade às pessoas moldarem e influenciarem as formas de poder cultural. No entanto, Dan Wielsch (2020) explica que no ambiente online a formação do sentido não é livre e não se dá por conta do processo de compreensão social porque a seletividade algorítmica causa uma aceitação da comunicação, guiando o comportamento dos usuários por protocolos de rede e comandando a formação de expectativas sociais, por isso ele afirma que “A tecnologia de rede é, por assim dizer, um meio de difusão e resultado a um só tempo.” (WIELSCH, 2020, p. 97).

No tocante ao terceiro valor da liberdade de expressão que é o crescimento e a disseminação do conhecimento, Balkin (2020) esclarece que não usa a expressão mercado de ideias porque para ele a competição de opiniões não é melhor forma de disseminação do conhecimento tendo em vista que, nas sociedades modernas, a disseminação e o desenvolvimento do conhecimento dependem de disciplinas, instituições, como escolas e universidades, e públicos relacionados a profissões. Para o mesmo autor, fontes plurais e antagônicas significam instituições diversas a produzir conhecimento e não vozes diferentes que discordam e disputam como numa competição de mercado.

Assim, ao transpor as discussões sobre a liberdade de expressão para o cenário da comunicação online, percebe-se que o argumento liberal da “competição do mercado” no sentido de propagar a verdade para descortinar as fake news pode ter o efeito inverso do desejado. Nesse sentido, Cass Sunstein (2010), argumenta que combater a desinformação com mais informação pode trazer ainda mais desconfiança e novamente fazer circular juízos e interpretações prejudiciais sobre fatos, ao invés de sedimentar a verdade e frear as falsidades.

Como relação aos valores da da liberdade de expressão de Balkin (2020), as fake news os subvertem na medida em que não se prestam a informar porque a natureza é de conteúdo desinformativo, também prejudicam a formação do sentido sendo manejadas para alterar a percepção da realidade, assim como não ajudam a promover o conhecimento tendo em vista que, sistematicamente, lançam dúvida, minando a confiança e a credibilidade das pessoas e das instituições que produzem conhecimento.

A liberdade de expressão é essencial para a formação da personalidade do indivíduo, assim como é para a democracia política, para a democracia cultural e para a disseminação do conhecimento. Os autores Ingo Wolfgang Sarlet e Ilton Robl Filho traduzem bem a dupla concepção do direito à liberdade de expressão.

Assim como a liberdade de expressão e manifestação do pensamento encontra um dos seus principais fundamentos (e objetivos) na dignidade da pessoa humana, naquilo em que diz respeito à autonomia e ao livre desenvolvimento da personalidade do indivíduo, essa liberdade também guarda relação, numa dimensão social e política, com as condições e a garantia da democracia e do pluralismo político, assegurando uma espécie de livre mercado das ideias e assumindo, neste sentido, a qualidade de um direito político com dimensão nitidamente transindividual, pois a liberdade de expressão e os seus respectivos limites operam essencialmente na esfera das relações de comunicação e da vida social. (SARLET; FILHO, 2016, p. 119)

A liberdade de expressão é essencial para formação do indivíduo porque a medida que revela opiniões, anseios, expectativas e sonhos se torna único. Na perspectiva democrática, a liberdade de expressão é um direito subjetivo do indivíduo que impõe uma abstenção estatal para que o cidadão não tenha receio de retaliações em razão das críticas ao poder público assim como também é um direito/dever de obter informações do Estado e contribuir para a vida em sociedade. Corroborando com esse entendimento:

Se a liberdade de expressão e comunicação, nos seus primórdios, estava ligada a dimensão individualista da manifestação livre do pensamento e da opinião, viabilizando a crítica política contra o *ancien régime*, a evolução daquela liberdade operada pelo direito/dever à informação, especialmente e corretamente informado; àquela dimensão individualista-liberal foi acrescida de outra dimensão de natureza coletiva: a de que a liberdade de expressão e comunicação contribui para a formação da opinião pública pluralista. (FARIAS, 2008, p.149).

Assim, a liberdade de expressão é direito negativo, de abstenção quanto a críticas, e um direito positivo do cidadão em que o Estado deve garantir meios de participação política. A participação informada auxilia na formação da vontade política baseada na pluralidade de opiniões. Diante disso, se mostra importante entender a liberdade de expressão não apenas em seu aspecto subjetivo mas também como mandamento jurídico-objetivo de resguardar a comunicação social, de viabilizar o exercício da cidadania com informações verdadeiras e de manutenção da pluralidade democrática, para impedir o efeito silenciador de minorias e para

evitar manipulações da opinião pública diante da assimetria de poder político-econômico do Estado, das empresas gestoras da comunicação online ou mesmo frente aos novos arranjos sociais como o das milícias digitais.

Nesse sentido, expressar-se livremente e autogovernar-se não se justifica à medida que provoca prejuízos e propaga inverdades sobre fatos como é o caso das fake news que atingem a honra, imagem, privacidade e intimidade alheias, ferindo a dignidade, disseminando dúvida sobre políticas públicas, divulgando teorias de conspiração política, fragilizando os elos de confiança intersubjetivos, promovendo a instabilidade das organizações sociais e pondo em risco instituições como a democracia e a saúde pública ou mesmo a vida de outras pessoas. Desse modo, as fake news representam um exercício abusivo da liberdade de expressão.

Diante disso, pode-se dizer que as fake news se apresentam como um contravalor da liberdade de expressão porque deformam a informação, induzem a compreensão social com design e técnicas de controle de comportamento, distorcem a opinião pública e não constroem conhecimento, causando até mesmo o silenciamento de minorias por impor-lhes o medo do isolamento social, e quando disseminadas por robôs, não há a menor chance para a construção de conhecimento ou intercâmbio de ideias porque as máquinas não são permeáveis ao diálogo nem são capazes de refletir aos contra-argumentos.

Evgeny Morozov (2020) ressalta que, na perspectiva da inteligência artificial, os algoritmos processam a prevalência das mensagens, em razão da quantidade cliques e compartilhamentos e assim, de modo frio, os algoritmos processam as fake news como algo genuíno, como se vê:

O modelo de negócios da Big Tech funciona de tal maneira que deixa de ser revelante se as mensagens disseminadas são verdadeiras ou falsas. Tudo que importa é pela escala de contatos online que é mais expandida que os vínculos sociais é se elas viralizam (ou seja, se geram números recordes de cliques ou curtidas), uma vez que é pela análise de nossos cliques e curtidas, depurados em retratos sintéticos de nossa personalidade, que essas empresas produzem seus enormes lucros. Verdade é o que gera mais visualizações. Sob a ótica das plataformas digitais, as fake News são apenas as notícias mais lucrativas. (MOROZOV, 2020, p. 11).

As empresas gestoras da comunicação online, ao atuarem como empresas privadas perseguindo finalidades lucrativas, manejam as fake news no ambiente de discussão para mantê-lo em movimento, fignando a atenção dos usuários, prologando o tempo de conexão e extraindo dados da navegação para direcionar publicidade. É a reação dos usuários diante de uma mensagem que potencializa o espectro da divulgação algorítmica, sem que seja considerada a veracidade ou não do conteúdo.

Ainda que exista muita informação relevante circulando online, inclusive oriunda de fontes oficiais e verificadas, as fake news é que assumem maior visibilidade no ambiente de discussão online porque a infraestrutura privada que faz a gestão dos fluxos comunicacionais persegue lucros e prioriza conteúdo economicamente mais interessante. Assim, ao selecionar e direcionar as informações, inclusive fake news, as empresas interferem na compreensão social sobre a realidade, manipulando a opinião pública e influenciando ou mesmo determinando as escolhas políticas e processos eleitorais.

Ao explicar que o ambiente online é fragmentado, Paulo Brasil Menezes (2020) argumenta que as informações não circulam livremente, sendo direcionadas ao usuário por análises algorítmicas com o uso de IA, segundo preferências do perfil, comprometendo a pluralidade democrática e tornando a comunicação enviesada e artificial.

Ocorre que, diante o volume de dados na internet, Balkin (2018) entende que os intermediadores da comunicação se tornaram essenciais no papel de curadores da comunicação online porque facilitam a participação pública na arte, na política e na cultura, organizam as conversas públicas para que as pessoas possam interagir e, por último, curam a opinião pública com feeds personalizados, reforçando normas de civilidade, segurança e comportamento nos termos de serviços e nas diretrizes da comunidade.

De acordo com Balkin (2018), as empresas de tecnologia, antes meros intermediárias dos fluxos comunicacionais, criaram uma burocracia interna com mecanismos de controle de conteúdo online à medida que se expandiam internacionalmente, treinando moderadores humanos e aperfeiçoando algoritmos para aplicar suas regras mais uniformemente. Assim, de acordo com o autor, quase por acidente, as empresas de tecnologia se tornaram empresas de mídia digital, curadores digitais.

Dito isso e reportando às consequências dessa metamorfose para o âmbito nacional, as empresas de tecnologia ao se tornarem empresas de mídia social e curadores de opinião passaram a se enquadrar como meios de comunicação social eletrônica e devem observância aos princípios insculpidos no art. 221 da Constituição Federal de 1988.

Art. 221. A produção e a programação das emissoras de rádio e televisão atenderão aos seguintes princípios:

I - preferência a finalidades educativas, artísticas, culturais e informativas;

II - promoção da cultura nacional e regional e estímulo à produção independente que objetive sua divulgação;

III - regionalização da produção cultural, artística e jornalística, conforme percentuais estabelecidos em lei;

IV - respeito aos valores éticos e sociais da pessoa e da família.

(...)

No capítulo 2 (tópico 2.3.1) foi demonstrado que a moderação de conteúdo - bloqueio, sombreamento, impulsionamento, exclusão - operada pelas empresas gestoras da comunicação online não tem se demonstrado uniforme, apesar de enumerar critérios e situações não há uma exposição clara sobre como suas normas são aplicadas. Para Evelyn Douek (2022) toda decisão de moderação envolve interesse comercial e ela explica que remover apenas conteúdo ilegal ou indesejável tornaria o ciberespaço inutilizável porque há muitos outros assuntos como assédio, spam, dentre outros discursos que não são ilegais, mas são do interesse dos legisladores e do público que sejam removidos, portanto, a temática a ser moderada é um rol em constante aperfeiçoamento.

De acordo com Jack Balkin (2018) o Estado, em virtude da incapacidade técnica, não regula mais diretamente os falantes, mas pressiona a infraestrutura das empresas privadas para a governança online. Segundo o autor, quanto maior a eficiência tecnológica, mais os Estados cooptam e coagem as operadoras da infraestrutura privada da internet a assumirem a função de regular e vigiar a fala.

Jack Balkin (2020) traça também uma distinção entre os tipos de serviços digitais e defende obrigações diferentes quanto a liberdade de expressão online. Segundo o autor, para serviços básicos de internet a resposta regulatória deve ser a não discriminação, para sistemas de pagamento a regulamentação deve ser sobre legalidade/ilegalidade das transações, fazendo a ressalva que transações legais também não devem discriminar então, a moderação de conteúdo deve ocorrer somente nas mídias sociais e nos mecanismos de pesquisa.

Desse modo, quanto mais eficientes em vigilância, policiamento e controle as empresas se tornam, mais essa infraestrutura privada é exigida pelos Estados. Nesse sentido, Balkin(2018) argumenta que as empresas de mídia social e mecanismos de pesquisa são pressionadas pelos governos para frear a disseminação de fake news e implementar políticas de discursos, no entanto, revela que há um antagonismo nas medidas porque, ao tempo que essas empresas estabelecem regramentos internos para excluir fake news, os incentivos econômicos estimulam a produção e a circulação de fake news.

Como se vê, o atual modelo de regulação do discurso digital é direcionado a infraestrutura da internet e não ao falante propriamente. Para Balkin (2018) essa cooperação e cooptação público-privado ocorre porque, para o Estado, é mais fácil regular e localizar o operador ou proprietário da infraestrutura do que identificar usuários que podem ser anônimos, estar além da fronteira ou nem mesmo serem humanos. O autor complementa ainda que a

infraestrutura privada é mais receptiva a pressão do Estado e têm interesse em cooperar para expandir seus negócios e alcançar mais lucros.

Desse modo, Estado e empresas de tecnologia interagem em regramentos e formas de controle da atividade online e, em virtude da importância na vida cotidiana, Nicolas Suzor (2018) defende que os termos e serviços, contratos firmados entre empresas e usuários se acoplam como pactos constitucionais na medida em que disciplinam como os espaços sociais compartilhados são constituídos e governados, fazendo a ressalva de que a relação que se estabelece entre usuários e empresas é diferente da soberania-cidadã.

Nesse sentido, ao analisar a teoria constitucional diante das tendências de digitalização, privatização e a globalização, Gunther Teubner (2017) constata que houve uma mudança de soberania do Estado-Nação para regimes transnacionais por isso ele defende que o controle da liberdade de expressão migrou do Estado-Nação para instituições da internet. Diante desse cenário, o autor argumenta que na sociedade global está surgindo uma multiplicidade de constituições civis fora da política institucionalizada, como subsistemas para regular diversas áreas representativas da sociedade tendo em vista a improbabilidade de uma Constituição global unitária.

Em virtude dessas mudanças provocadas pela internet, em que o espaço público era restrito às fronteiras dos países e expandiu para domínios privados desterritorializados, Gunther Teubner (2017) entende que os direitos constitucionais podem proteger não apenas direitos individuais, mas também instituições sociais perante os meios de comunicação expansionistas. Para o autor, a contextualização de direitos constitucionais deve se concentrar nos riscos para as instituições sociais autônomas, ou seja, no perigo que o próprio código digital produziu. Desse modo, a governança online e os arranjos sociais que a internet propiciou, como as milícias digitais que ameaçam a democracia e escapam aos controles públicos, são fatores a serem considerados nos conflitos estruturais da sociedade conectada.

No Brasil, a atividade estatal é limitada por direitos e garantias previstos constitucionalmente em face ao Estado. No âmbito privado, a relação empresa-usuário, vigora a cláusula geral da legalidade, segundo o qual o que não está proibido é permitido, prevista no art. 5º, II, da CF/88. Desse modo, as empresas podem instituir regras a serem obedecidas dentro de seus domínios, desde que não sejam expressamente vedadas pelo ordenamento jurídico nacional.

Assim, as gestoras da comunicação online não podem impor restrições que causem danos ou ofensas ou que repercutam negativamente para a coletividade, sendo no entanto admitidas políticas internas e termos de uso que otimizem as atividades, em observância aos fundamentos do Estado Constitucional. A questão é que a lógica dos fluxos comunicacionais é baseada no lucro e, desse modo a circulação da informação no ambiente de discussão online, se distingue da forma como a informação transita num regime democrático.

As práticas opacas de moderação e modulação de conteúdo visam lucros e desse modo as informações são selecionadas e dirigidas para públicos específicos segundo características catalogadas no perfil dos usuários, compartimentando a informação que deveria circular em igualdade de condições para todos, formando segmentos populacionais, com audiências isoladas e sem contato com posições divergentes.

Num país democrático, as informações deveriam ser dirigidas sem distinções dos grupos populacionais, ou ao menos com a advertência de que o fluxo não é uniformemente distribuído para todos, de modo que torne claro que as informações são fragmentadas, tornando transparente os critérios que motivam os direcionamentos e dando oportunidade para os usuários alterarem dados no intuito de democratizar o que lhes é enviado.

Informações selecionadas para grupos populacionais segmentados criam ambientes fragmentados e, assim, as análises algorítmicas com o uso de IA fabricam artificialmente mundos particulares onde não deveria haver divisões, destruindo a realidade compartilhada. Portanto, a resposta é que o direcionamento de informações online desvirtua o direito de informar e de ser informado. Essas práticas também criam obstáculos a comunicação social que não atinge o fim a que se propõe que é o de contribuir no debate entre posições divergentes e convergentes para a formação de posições e construção de decisões políticas conciliadas.

Na medida em que análises algorítmicas fazem escolhas de conteúdo, priorizam uns assuntos em detrimento de outros, e estruturam a comunicação ao organizá-la no feed de notícias, manejam o alcance e a visibilidade das manifestações segundo interesses comerciais, sem que as informações sejam distribuídas igualmente e, assim, a liberdade de expressão deixa de atender a democracia política porque os cidadãos não recebem informações balanceadas de modo que a opinião pública é distorcida e não está em sintonia com a realidade.

Do mesmo modo, a liberdade de expressão passa a não corresponder a democracia cultural porque a formação de sentido já chega pronta, não há interação entre os cidadãos e até mesmo a produção e a distribuição de conhecimento é prejudicada porque não está embasada e

fundamentada numa análise reflexiva, racional e científica, tendo em vista que o conteúdo online costuma circular e ser compartilhado por fatores reativos, emocionais e econômicos.

Assim, no modelo de negócios online, as informações são direcionadas ao cidadão do mesmo modo que os produtos e serviços são oferecidos ao consumidor, considerando renda, histórico de compras, interesse de pesquisas em sites de busca dentre outros. A dinâmica de oferecer e direcionar conteúdo segundo preferências do perfil acaba por não ofertar ao público a pluralidade de ideias, pensamentos e argumentos que existem no ambiente democrático e isso, conforme Vânia Baldi (2018), reforça concepções e radicaliza opiniões.

Bobbio (2019, p. 92) entende que “pode muito bem existir um Estado democrático numa sociedade que a maior parte das instituições – da família à escola, da empresa à gestão de serviços públicos – não são governados democraticamente”, mas o autor faz a seguinte indagação: “a democracia política foi e é até agora necessária para que um povo não caia sob um regime despótico; mas é também suficiente?” Essa pergunta já traz em si uma inquietação, a de que nem tudo está bem. Some-se a indagação de Bobbio (2019), a conclusão de Kate Klonick (2017), segundo a qual a governança em rede não é democrática. No contexto tecnológico, a pergunta de Bobbio pode ser contextualizada assim: A democracia resiste ao modelo de negócios da governança online que modula e modera os fluxos comunicacionais segundo finalidades lucrativas?

Como explicitado acima, a liberdade de expressão online não tem servido à democracia política, representando até mesmo um risco para si própria, tendo em vista o poder que as empresas gestoras dos fluxos comunicacionais online têm para limitar a liberdade de expressão a ponto de impedir a participação, converter a liberdade de expressão em instrumento de agressão e radicalização e personalizar o conteúdo distribuído a cada usuário de acordo as preferências do perfil, estruturando as informações de modo a dividir as pessoas que tem opiniões divergentes, fabricando realidades particulares e prejudicando o debate democrático e plural (CALLEJÓN, 2022).

Ao analisar a crise do constitucionalismo, Callejón (2018) argumenta dois fatores que a motivaram, sendo um externo e outro interno. O fator externo está relacionado à globalização em que nem o Estado nem as organizações supranacionais têm conseguido viabilizar direitos ou promover consensos democráticos em virtude da capacidade de ação limitada. Como fator interno, refere-se aos processos comunicativos, em que o autor menciona expressamente as redes sociais, por promoverem uma mudança no jogo de forças que antes se resolvia pela pluralidade democrática.

As reflexões de Bobbio e Callejón tem como ponto em comum a fragilidade da democracia. Bobbio (2019) se refere a democracia política e a convivência com instituições não democráticas. Callejón (2018) trata da crise do constitucionalismo em que a canalização dos conflitos sociais segundo os processos democráticos, com a internalização do pluralismo e composição dos interesses divergentes para a geração de acordos, se mostra insuficiente porque as práticas online, ao priorizar fake news, destroem a realidade social compartilhada, promovendo instabilidade e insegurança.

A democracia tem convivido com o modelo de negócios da comunicação online que Klonick (2017) considera como uma governança não democrática, Balkin (2018) julga a governança online como autoritária e Gunther Teubner (2017) a considera com tendências expansionistas. Assim, o fluxo de informações tem deturpado os processos de composição das divergências e pluralismos em decisões políticas convergentes tendo em vista o crescimento de radicalismos e polarizações fomentados por fake news, que se tornam ainda mais perigosas quando disseminadas por milícias digitais.

Desse modo, as fake news têm se mostrado um fator de risco para a democracia. Callejón (2018) trata da progressiva desintegração do espaço público, essencial para a formação da opinião pública. Para o autor (CALLEJÓN, 2018), os fluxos comunicacionais online, especialmente nas redes sociais, causam desequilíbrio no contexto democrático porque fragmentam e polarizam as pessoas. Também tratando sobre desequilíbrio no contexto democrático, Bobbio (2019) defende que existe um limite do dissenso que as democracias suportam. Em termos gerais, os autores mencionam o nível de discordância e divergência que o regime democrático consegue dissolver através dos processos democráticos no sentido de manter a coesão social.

A estruturação da informação no ambiente de discussão online, que privilegia fake news para manter a agitação online e que seleciona conteúdo de acordo com as preferências do perfil restringindo a pluralidade que toca cada usuário, divide as pessoas em grupos opostos e as distancia. Desse modo, a forma como o fluxo comunicacional é organizado no ambiente de discussão online não se mostra favorável ao diálogo colaborativo e pode mesmo inviabilizar os consensos democráticos. Diante disso, o estudo considera as fake news disseminadas por milícias digitais para além dos limites do dissenso que Bobbio (2019) afirma existir.

Como há uma articulação coordenada em rede, o ataque de fake news pelas milícias digitais foi considerado pela pesquisa como subversivo às regras do jogo democrático justamente por agir e estimular a conflituosidade e polarização em escala ao invés do consenso

e do entendimento. Assim, o disparo de fake news pelas milícias supera os limites do dissenso na medida que inviabiliza o processo de internalização das diferenças para a construção uma decisão conciliada. O dissenso é da essência da democracia em sociedades de massa, mas apesar disso, há um limite para os dissensos que as democracias suportam e são capazes de processar, absorver e converter em decisões conciliadas. Ultrapassado esse limite, a democracia não funciona como um sistema para compor opiniões e posições divergentes.

No ambiente online o argumento contrário às fake news e que as descortina é racional e não tem o apelo às emoções nem a mesma visibilidade online porque o modelo de negócio é projetado para priorizar fake news por serem economicamente mais interessantes já que afetam os sentidos e provocam maior agitação e engajamento. Desse modo, no ambiente online não parece haver uma força/argumento, em igualdade de condições, capaz de contrapor e disputar com as fake news disseminadas por milícias digitais, de modo a propiciar em ambiente online equilibrado, com informações balanceadas. Desse modo, o comportamento humano ainda que racional e reflexivo não seria suficiente para frear fake news online disseminadas por milícias porque além de acionarem os mecanismos algorítmicos de impulsionamento ainda contam com o uso de robôs que incrementam artificialmente a disseminação.

A comunicação online tem desvirtuado os valores da liberdade de expressão e isso tem testado a resiliência do regime democrático, de modo que o pluralismo nas redes se tornou polarização, os extremismos e minorias têm mais atenção e espaço que as posições majoritárias da realidade física. A forma como a comunicação online é estruturada para promover agitação e os lucros faz crer que a atividade miliciana tem uma aceitação popular muito maior imbuindo o receio de isolamento social e isso altera a leitura da realidade compartilhada trazendo a falsa sensação de crise e iminente reviravolta.

Assim, de acordo com Francisco Balaguer Callejón (2022a), apesar do modelo de negócios não ter uma preferência ou um interesse político específico nem pretender impor uma narrativa concreta, as empresas que fazem a gestão dos fluxos comunicacionais acabam privilegiando fake news, não pelo conteúdo, mas pela agitação que promovem, e assim convertem a liberdade de expressão em instrumento de agressão e radicalização.

Desse modo, na disseminação de fake news por milícias digitais, a liberdade de expressão passa a ser instrumentalizada e deixa de ser um direito que possibilita o desenvolvimento pessoal ou os debates democráticos. Nesse sentido, o pluralismo é substituído por uma hierarquização algorítmica que organiza e estrutura a diversidade social escolhendo e

selecionando conteúdo de acordo com finalidades lucrativas e não com base em interesses sociais ou tendências majoritárias.

Ao estabelecer como preferência conteúdo promove que agitação e prende a atenção dos usuários, as posições extremistas e fake news são privilegiadas dentro do fluxo comunicacional. Isso explica por que o dissenso da realidade física é retratado no ambiente online com muito mais poder do que verdadeiramente possui e a construção de uma falso cenário conduz a uma leitura e interpretação da realidade compartilhada segundo uma tendência de instabilidade e insegurança.

Bobbio (2019, p. 100) entende que o pluralismo democrático “deve poder contar com um controle recíproco entre os grupos que representam interesses diversos, os quais se exprimem por sua vez através de diversos movimentos políticos que lutam entre si pela conquista temporária e pacífica do poder”. Para o autor, dentro do contexto democrático há disputa de opiniões divergentes que interagem na formação da opinião pública e na tomada de posição, de modo que os consensos, numa democracia representativa, são tomados por maioria e, apesar do dissenso real ser uma característica da democracia moderna, as forças no jogo democrático devem atuar dentro de limites, sem abusos, portanto, o jogo democrático não é um vale-tudo.

As milícias digitais usam a infraestrutura em rede e a lógica dos algoritmos para promover maciçamente uma posição que lhes favorece, disseminando fake news. Assim, apesar de Callejón (2022a) defender que os gestores da comunicação não têm um interesse político específico, as análises algorítmicas com o uso de IA priorizam fake news e as milícias digitais se valem disso. O comportamento articulado online das milícias digitais potencializa a distribuição das fake news na medida em que atença os algoritmos à engendrarem favoravelmente, ou seja, a ação externa promovida pelo ataque de milícias digitais aciona os mecanismos internos de impulsionamento de conteúdo. É uma conjunção tecnológica que torna o disparo de fake news por milícias digitais uma força sobre-humana, destrutiva, comparável a uma doença, como uma disseminação viral.

Apesar do ambiente de discussão online ser um espaço privado influencia a democracia porque o ambiente de discussão online também funciona como um termômetro para os parlamentares que muitas vezes conduzem sua atividade legislativa segundo a expectativa dos eleitores que é imediatista e sem uma diretiva a longo prazo. A democracia política é, segundo Balkin (2020), baseada em participação, receptividade e público informado. Ocorre que a democracia política tem sido ignorada pela gestão autoritária em virtude de as práticas

operacionais da comunicação online prestarem informações superficiais e incompletas, além de serem fragmentadas para diferentes grupos populacionais e onde nem todos tem as mesmas oportunidades de serem ouvidos.

Os meios de produção de informações de acordo com padrões profissionais de jornalismo estão sendo minados por conta do modelo de negócios de online que monopoliza o mercado publicitário, fonte de receitas das organizações de imprensa. Considerando que as pessoas, em maioria, se informam no meio virtual, a atuação dos gestores dos fluxos comunicacionais online têm posto em risco a democracia política tornando a inquietação e a pergunta de Bobbio (2019), sobre a democracia e seus processos como garantias que impedem a ascensão regimes despóticos, cada vez mais viva e atual tendo em vista que as práticas online minam o diálogo colaborativo e, ao prestigiarem as fake news, acabam por diminuir a circulação de conteúdo produzido segundo padrões éticos e profissionais e princípios informativos.

Os usuários, são também, em regra, cidadãos, no entanto, Estado e empresas não se confundem, nem perseguem os mesmos fins por isso é importante ressaltar e diferenciar a relação entre empresa-usuário da relação Estado-cidadão. A relação Estado-cidadão persegue fins públicos e é definida pela CF/88 que estabelece os direitos dos cidadãos diante do Estado e fixa os limites da atuação estatal. De outro modo, a comunicação online é operada segundo um modelo de negócios que persegue lucros, sendo a relação empresas-usuários orientada pelo direito privado, com cláusulas contratuais firmadas entre as partes e por regras internas.

Ocorre que as empresas que fazem a gestão dos fluxos comunicacionais online estão inseridas no quadro da comunicação social eletrônica e por essa razão exercem funções públicas. Para Humberto Martins (2018, p.10), a Constituição Federal de 1988 ampliou “o conceito de público para além das fronteiras do Estado. O público foi afirmado como imerso na sociedade.” Traz ainda que “as tecnologias da informação e da comunicação desempenham papel central para esta concretização de direitos, como é muito claro ao tratarmos de transparência.”. Segundo ele, o processo tecnológico de informação e comunicação social é um catalisador de anseios sociais, expectativas constitucionais e de transparência.

No entanto, a canalização das conflitos e das divergências por meio dos processos democráticos, através da internalização dos dissensos para a formação de consensos, está comprometida pelo modelo de negócios da comunicação online que faz a gestão dos fluxos em busca de lucros. Ao fragmentar informações e notícias segundo preferências do perfil, divide os cidadãos em compartimentos populacionais distintos e, ao omitir toda a diversidade que não

lhes foi direcionada, radicaliza opiniões e convicções prévias, distanciando as pessoas em pólos opostos e dificultando decisões compartilhadas.

Assim, o dissenso no ambiente de discussão online tem uma projeção muito maior do que na realidade física e, como o virtual influencia o real, as tomadas de posição são dificultadas pelas práticas online que promovem a destruição da realidade compartilhada. A distribuição dos fluxos comunicacionais online, tendo em vista perseguirem lucros, não são um reflexo fidedigno do pluralismo político nem da diversidade social. A internet não tem sido o espelho dos anseios sociais, o modelo de negócio mais se parece com uma lente de aumento para minorias e extremismos.

Se a internet, efervescente e de intensa interação entre os usuários, abriu espaço para temas que não são rentáveis no âmbito de uma indústria cultural de massa (LOVELUCK, 2018), trouxe também mais visibilidade aos dissensos. Não se conhece ao certo os efeitos concretos da influência online nem os limites toleráveis do dissenso numa democracia, no entanto, pode-se argumentar que, por estar está ocorrendo uma involução da comunicação social, as fake news têm papel importante na depreciação dos debates democráticos porque alimentam guerras de narrativas, radicalizam opiniões e dividem as pessoas.

Desse modo, na gestão dos fluxos online devem incidir os princípios insculpidos constitucionais referentes a comunicação social assim como as regras que disciplinam a relação usuário-empresa devem estar dentro de limites constitucionais, ou seja, devem considerar os anseios de uma comunicação social confiável e de contratos com cláusulas transparentes e menos obscuras, especialmente, porque essas regras implicam riscos a democracia tendo em vista que os milhões de contratos individuais são comunidades gigantes influenciadas pela estruturação e organização do fluxo comunicacional online.

Nesse ponto, apesar da governança das redes não ser democrática deve estar dentro de limites constitucionais. Isso significa que a gestão das redes conduzida por empresas privadas não precisa conter práticas democráticas na forma de organização interna (como eleição dos representantes, mandato por prazo determinado, decisão por maioria da comunidade), mas tem de estar submetida aos ditames constitucionais brasileiros. Luis Roberto Barroso sintetiza bem a diferença entre democracia e constitucionalismo:

O constitucionalismo democrático foi a ideologia vitoriosa no século XX. Nele se condensam as promessas da modernidade: poder limitado, dignidade da pessoa humana, direitos fundamentais e justiça material, tolerância, e o outro e - quem sabe? - até felicidade.

Nesse arranjo institucional se fundiram duas ideais que percorrem trajetória diferentes: o constitucionalismo, herdeiro da tradição liberal que remonta ao final do século XVII, expressa a ideia de poder limitado pelo direito e respeito aos direitos fundamentais. A democracia, por sua vez, traduz a ideia de soberania popular, de governo da maioria, que somente se consolida, verdadeiramente, ao longo do século XX, com a consagração do sufrágio universal e o fim das restrições a participação política decorrentes do nível de riqueza, do sexo ou da raça. (BARROSO, 2021, p. 68).

O desafio é fazer caber os parâmetros de governança online nos limites constitucionais. Nesse aspecto, para Francisco Callejón (2022a) quem não tem estrutura democrática não pode contribuir com a democracia. Segundo Teubner (2017), ainda que a comunicação através da internet tenha alguma regulação, inclusive constitucional, são os constrangimentos eletrônicos que regulam os comportamentos online porque o código transforma a ordem normativa no ciberespaço. Assim, diante do modelo de negócio e dos incentivos econômicos, não se pode esperar que o próprio código construa uma constituição digital transnacional que efetivamente limite seu próprio poder e proteja direitos.

Gunther Teubner (2017) aponta três questões que dificultam a construção de uma constituição digital pelo próprio “código”. Primeiro, como o código é autoimposto, há uma fusão entre as funções de criação, execução e aplicação das normas internas, em arrepio a legalidade. Em segundo, a tríade de: regulação da conduta, construção de expectativas e resolução de conflitos é reduzida a regulação eletrônica da conduta já que as práticas de governança são opacas, não há um sistema de controle interno e nem sempre se guiam de acordo as expectativas dos usuários. Por último, segundo ele, a normatividade digital é binária: excluir ou incluir, não há espaço para interpretação, proporcionalidade e contextualização. Desse modo, o autor defende que a constituição digital transnacional será fruto de atividade dos tribunais a ser construída por meio de revisão das cláusulas de contratos digitais padrão.

Ocorre que, apesar de Gunther Teubner (2017) não acreditar em soluções globais a respeito do fluxo de comunicações online por parte das empresas gestoras, o Estado não dispõe dos meios tecnológicos, então a pressão e cooptação da regulação proveniente de grandes blocos entre os países deve aumentar no intuito das empresas criarem obrigações de transparência e eficiência e, desse modo, mecanismos de controle de fluxos sejam incentivados e aperfeiçoados.

Nesse sentido, Balkin (2018) explica que as possibilidades técnicas e jurídicas das empresas privadas para regular o discurso são muito maiores do que as do Estado, no entanto, ele argumenta que como as empresas selecionam e personalizam as informações, elegendo assuntos preferenciais e tomando decisões sobre conteúdo, precisam ser confiáveis. O autor

argumenta que os curadores digitais, mídias sociais e mecanismos de pesquisa, são vistos como facilitadores da participação cultural e em razão disso os usuários esperam e dependem da ajuda dessas empresas, por isso ele entende que existem obrigações sociais e morais perante o público e defende que a intermediação da comunicação requer um ambiente confiável e guiado por normas profissionais.

Ocorre que, segundo Jonathan Albright (2017), nas redes a confiança é aferida pela codificação dos sentimentos e a monetização da atenção e não pelo conteúdo, por isso as interações entre os usuários são projetadas e manipuladas por emoções e desejos como o botão de curtidas e emojis e pelos lucros. Desse modo há um conflito entre as expectativas dos usuários, confiabilidade e transparência, e o modelo de negócios online das empresas que visa os lucros, repercutindo em choques com os princípios da comunicação social e com os anseios do público.

Para Balkin (2018), a moderação de conteúdo feita por empresas privadas pode implicar maiores riscos para a liberdade de expressão porque, quando em dúvida se o conteúdo é ou não ilegal, podem optar por remover o discurso, agindo por excesso de cautela a fim de evitar multas e sanções, causando o que ele chama de censura colateral. Para ele (BALKIN, 2018) aumentam também as chances de restrição digital prévia em que decisões de bloqueio de fala ou de perfil não são transparentes e ocorrem sem que tenha havido ordem judicial, ficando a liberdade de expressão digital a mercê de uma decisão interna sem nenhum direito processual como exposição da motivos ou prazo para responder uma reclamação.

Desse modo, para Balkin (2018), como as empresas removem, filtram e bloqueiam conteúdo, ao usuários têm direito ao devido processo, direito de clareza e aviso sobre as regras internas das empresas, explicação e oportunidade de ser ouvido antes da remoção da postagem ou, na impossibilidade, o direito de explicação e revisão *post fact*, como garantias processuais mínimas do Estado de Direito.

Em estudo posterior, Balkin (2020, p. 06) afirma que os intermediadores da comunicação online, empresas privadas, estão livres para impor limitações ao discurso como normas de civilidade, segurança e comportamento. Nesse mesmo ensaio, o autor (2020) também reconheceu que as mídias do século XX, normalmente a mídia de massa, limitavam a fala muito mais do que a lei exigia estabelecendo o conteúdo permissível baseado em interesses e valores esperados de sua audiência. Assim, até mesmo Jack Balkin, defensor da liberdade de expressão, reconhece que nas mídias sempre houve limitações além das legais.

Numa percepção mais pragmática sobre a comunicação online, Evelyn Douek (2022) sustenta que exigências processuais excessivas não se compatibilizam com moderação de conteúdo online, privatizada, rápida e em larga escala sob o risco de tornar a moderação lenta, cara e impraticável por empresas menores, restringindo o acesso ao mercado e dificultando a inovação. Para a autora (DOUEK, 2022) tentar corrigir os erros de moderação com as garantias processuais aplicando a doutrina do constitucionalismo digital, por exemplo, acaba por sugerir reformas contraproducentes na medida que terão por consequência um sistema recursal privatizado abarrotado como o Poder Judiciário e cujo resultado, a decisão final, mesmo tendo sido ponderado e discutido enfrentará desacordos porque não existe uma resposta certa e definitiva.

Como exemplo, Evelyn Douek (2022) fala sobre a exclusão de vídeos que transmitiam o atentado na mesquita de Christchurch na Nova Zelândia em 2019 em que 51 pessoas foram mortas, tendo o atirador planejado toda ação para ser transmitida ao vivo pelo Facebook. Sobre esse ataque, Katie Klonick (2019) complementa que, para retirar o conteúdo de suas redes, o Facebook criou uma identidade digital das imagens e do som e excluiu 1,5 milhões de cópias, sendo que 1,2 milhões foram no momento do upload. Nesse sentido, Evelyn Douek (2022, p. 23), afirma que para tomar decisões oportunas, no sentido de rápidas e eficazes, as falhas são inevitáveis, como as que ocorreram com a filtragem do massacre que acabou por excluir conscientemente reportagens com partes do vídeo. As ferramentas automatizadas são insensíveis as nuances, desse modo, aceitar e escolher as falhas é uma consideração do design, especialmente se comparadas às compensações em escala.

Assim, ela (DOUEK, 2022) sustenta que seria irrazoável ter de comunicar cada um dos usuários que teve conteúdo excluído mediante filtragem em massa e por isso afirma que aplicar o modelo de moderação de conteúdo caso a caso na escala global é complicado porque perpassa pela necessidade de fornecer avisos de remoção de conteúdo com fundamentação, propiciar um meio para apelar e o direito a um revisor humano independente.

Em virtude das várias possibilidades de moderação de conteúdo, Evelyn Douek (2022) defende uma compreensão sistêmica do discurso considerando que, com a tecnologia online, a liberdade de expressão adquiriu alcance e velocidade incompatíveis com a adjudicação offline e que, em virtude da escala e do design, as medidas de controle online tem funcionalidades algorítmicas que se operam de maneira automatizada.

Para Evelyn Douek (2022), uma visão sistêmica sobre a moderação de conteúdo implica uma responsabilidade geral das empresas de mídia. Ela argumenta que, paulatinamente, esse

tipo de controle tem sido feito pelas empresas, e que também foi verificado pela pesquisa no tópico 2.3.1 na medida em que cada rede amplia o espectro de assuntos moderados como demonstrado em capítulo 2. Eleger a análise no modelo singular, caso a caso, como um modelo padrão de moderação negligencia as muitas outras formas de tomada de decisão de conteúdo, no entanto, é preciso ser a tolerante ao risco de falhas especialmente porque nenhum modelo é perfeito.

Como visto, a moderação de conteúdo sempre traz um efeito inibitório a liberdade de expressão e é justificável na medida em que se reconhece que a liberdade de expressão não é direito absoluto. A comunicação social sempre operou com restrições à liberdade de expressão, seja porque nem todos tinham acesso às mídias tradicionais já que era priorizado para jornalistas profissionais ou porque o controle editorial prévio estabelecia quais os assuntos seriam abordados e como seriam conduzidos. Nem todos e nem tudo eram permitidos. Além disso, a responsabilidade de indenizar ou conceder direito de resposta também eram circunstâncias limitadoras dos discursos, que promoviam uma reflexão sobre as consequências das publicações e continham a manifestações por prudência, sopesando consequências como deveres e responsabilidades.

Como na comunicação online a checagem do conteúdo não costuma ser prévia e em virtude das proporções em escala do universo virtual é preciso que sejam desenvolvidos filtros automatizados para contingenciamento de fake news, especialmente as disseminadas por milícias digitais, cuja propagação viral articulada, rápida e maciça, torna a contenção dos efeitos e a reparação dos danos ainda mais difíceis. Desse modo, a forma de contingenciamento a ser manejada teria de examinar as características da divulgação online como o conteúdo, o contexto, a articulação em rede e o uso de robôs para avaliar qual o melhor modelo de controle.

Evelyn Douek (2022) defende que haja uma tolerância ao risco na moderação sistêmica, operada de maneira automatizada porque nenhum modelo é infalível e existem outros benefícios desse modelo a serem considerados como a exclusão da postagem pela filtragem e a rapidez com que se opera. A autora também sustenta que os recursos devem ser priorizados não apenas com base no tempo da decisão, mas também considerando a gravidade das consequências, o potencial de danos e o tipo de decisão. Segundo a autora, remoção de conta requer proteção maior do que exclusões de conteúdo e, desse modo, o controle e reclamação do usuário devem abordar outras situações que não apenas a cronologia, mas a gravidade dos danos e as características da divulgação.

A escala da comunicação online e a velocidade de propagação torna a moderação de conteúdo online baseada em casos individualizados muito simplista diante das possibilidades tecnológicas. Ainda segundo Evelyn Douek (2022), as regras processuais devem ser concebidas como um sistema a luz do objetivo geral dos programas então aplicar os mesmos procedimentos de adjudicação offline para a governança online acaba por não considerar as interferências governamentais, as filtragens algorítmicas e o controle do design na moderação online.

Diante disso o presente estudo aborda a possibilidade de filtragens algorítmicas de fake news disseminadas por milícias digitais sob a perspectiva dos fundamentos constitucionais. A Constituição vigente foi promulgada após a ditadura militar e assume a defesa da liberdade de expressão como ponto central para a proteção do regime democrático (MACEDO JR, 2020). As limitações da liberdade de expressão pelo poder estatal foram elencadas no Título II – Dos Direitos e Garantias Fundamentais, Capítulo I – Direitos e Deveres Individuais e Coletivos, nos primeiros artigos do texto constitucional, logo em seguida a afirmação dos direitos.

Assim, de acordo com a técnica de redação utilizada, primeiro há o reconhecimento do direito a manifestação do pensamento, em seguida a vedação ao anonimato (art. 5º, IV, CF/88) e o direito de resposta (art. 5º, V, CF/88). Do mesmo modo, no art. 5º, IX, CF/88, primeiro se garante a liberdade de expressão intelectual, artística, científica e de comunicação para então mencionar que essas atividades são independentes de censura ou licença. Essa disposição também se repete no art. 5º, X, CF/88 que primeiro garante a inviolabilidade da intimidade, da vida privada, da honra e da imagem para, em seguida prescrever, a indenização por dano moral e material posterior como medida reparadora.

Art. 5º Todos são iguais perante a lei, sem distinção de qualquer natureza, garantindo-se aos brasileiros e aos estrangeiros residentes no País a inviolabilidade do direito à vida, à liberdade, à igualdade, à segurança e à propriedade, nos termos seguintes:

I - homens e mulheres são iguais em direitos e obrigações, nos termos desta Constituição;

II - ninguém será obrigado a fazer ou deixar de fazer alguma coisa senão em virtude de lei;

(...)

IV - é livre a manifestação do pensamento, sendo vedado o anonimato;

V - é assegurado o direito de resposta, proporcional ao agravo, além da indenização por dano material, moral ou à imagem;

(...)

IX - é livre a expressão da atividade intelectual, artística, científica e de comunicação, independentemente de censura ou licença;

X - são invioláveis a intimidade, a vida privada, a honra e a imagem das pessoas, assegurado o direito a indenização pelo dano material ou moral decorrente de sua violação;

(...)

XIV - é assegurado a todos o acesso à informação e resguardado o sigilo da fonte, quando necessário ao exercício profissional;

As limitações estabelecidas são garantias face ao Estado no intuito de preservar o ambiente democrático e o desenvolvimento pessoal. É natural que a Constituição Federal não tenha fixado restrições à liberdade de expressão no ambiente online porque a comunicação social eletrônica online era muito incipiente quando de sua promulgação, só vindo a ser reconhecida e acrescentada ao texto constitucional em 2002 com a promulgação da emenda constitucional 36.

Nesse cenário de grandes transformações induzidas pela Internet nos últimos 20 anos, as dinâmicas sociais foram fortemente impactadas assim como os direitos, exigindo adaptação para a nova realidade e é justamente esse o ponto de partida da pesquisa: a utilização da tecnologia de filtragem algorítmica para exclusão de fake news disseminadas por milícias digitais é constitucional? Sobre a releitura das situações e adaptações do direito em face de problemas e oportunidades mutáveis, Holmes e Sustain expõem:

Todos os direitos jurídicos – no direito constitucional como no direito privado – surgiram originalmente como respostas práticas a problemas concretos. Esta é umas razões pelas quais variam de época para época e de jurisdição para jurisdição. Na qualidade de instrumentos forjados para servir a interesses humanos e concepções morais mutáveis, são reiteradamente remodelados ou reespecificados por novas leis e decisões judiciais. Os direitos também se transformam por que os obstáculos ao bem-estar humano – os problemas que os direitos visam mitigar ou superar – mudam pari passu com a tecnologia, a economia, a demografia, os papéis profissionais, os estilos de vida e muitos outros fatores. (HOLMES E SUSTEIN, 2019, p. 189)

Assim, por se tratar de domínios particulares se reconhece a moderação e modulação de conteúdo não somente como um direito dessas empresas de atuarem em seus espaços privados, mas também como obrigação em razão dos riscos provocados por suas atividades lucrativas, de modo que seja útil à democracia e às vítimas, além de um canal para denúncias e reclamações para denúncia e exclusão de conteúdo, a supervisão das redes para evitar circulação de fake news ou material degradante, violento e criminoso. Nesse sentido, Evelyn Douek (2022) argumenta que a moderação de conteúdo é também do interesse do modelo de negócios, para evitar que os usuários se afastem.

No entanto, existem críticas às soluções implementadas pelas próprias empresas que impulsionam, incentivam ou estimulam fake news. O fato de privatizar a solução do problema às mesmas empresas que criaram o modelo de negócios da distribuição da informação aumenta a dependência da comunicação social eletrônica online em relação à infraestrutura privatizada. Nesse sentido, Evgeny Morozov (2020) critica a monopolização de soluções algorítmicas pelas Big Tech porque resulta em mais poder para essas empresas. Para ele, o falso verniz de objetividade dos números a serem apresentados como resultado das filtragens pode causar mais danos e ocasionar mais injustiças, se não houver uma preocupação com os dados usados para

fazer essas avaliações. O autor remete aos riscos de análises com vieses discriminatórios em escala como filtragens algorítmicas de grupos minoritários.

Nesse sentido, as filtragens algorítmicas para exclusão de fake news disseminadas por milícias digitais têm de ser imparciais, ideologicamente desinteressadas, sob pena de causar desequilíbrios democráticos e, ao invés de servirem como meios de correção e restauro do ambiente de debate e de ordem informacional, as filtragens serviriam a usuários economicamente privilegiados que poderiam pagar por esses serviços em rede no intuito de remover conteúdo que lhes seja desfavorável. Portanto, as filtragens não podem ser um serviço ou meramente um direito ou liberalidade das empresas gestoras dos fluxos comunicacionais, mas obrigações de prevenção e reparação diante dos riscos e dos danos a terceiros decorrentes suas atividades lucrativas.

A solução sobre o fluxo comunicacional a ser resolvido sob a perspectiva da educação digital, num trabalho de amadurecimento de cidadania no meio virtual, é uma tarefa hercúlea e, apesar de ser relevante política pública instruir as pessoas a checar informações antes de repassar, como uma conduta cívica consciente, não traz uma solução à matriz do problema, que é o modelo de negócios que remunera o conteúdo de acordo com o engajamento, comentários, compartilhamentos e curtidas. Nesse sentido, Jonathan Albright (2017), revela que as evidências baseadas em fatos deixaram de ser relevantes para parte da população e, por essa razão, a batalha das fake news não pode ser vencida apenas considerando o conteúdo porque a objetividade das notícias não é mais o que capta a audiência para avaliar verdade, justiça e precisão.

Há um desequilíbrio entre a capacidade humana de racionalizar um conteúdo comparada aos mecanismos artificiais de processamento de inteligência artificial que rapidamente selecionam e direcionam conteúdo. Nas palavras de Evgeny Morozov (2020, p. 177) fica mais compreensível: “[p]regar o autocontrole diante do extrativismo de dados é como pregar o empreendedorismo diante da destruição causada pelo neoliberalismo: é uma maneira de reduzir um problema coletivo e político ao nível individual, adequado ao consumo.” Isso significa que o fluxo de informações online é matematicamente tendencioso por ser uma consequência dos incentivos financeiros, e o combate às fake news não deve ser enfrentado unicamente pelos usuários de modo particularizado com reclamações individuais ou mediante um longo processo de amadurecimento cívico de milhões de pessoas.

Enquanto não se efetivam medidas que impedem a circulação, as fake news continuam a produzir efeitos seja ferindo a dignidade humana ou atingindo bens coletivos juridicamente

relevantes como a democracia. Os expedientes previstos na CF/88 como indenização e direito de resposta não promovem a justa reparação porque os danos online assumem proporções difíceis de prever e quantificar pela amplitude de propagação, além disso, a trajetória de um conteúdo online é imprevisível, razão pela qual o direito de resposta não atinge o memo público que as fake news.

Para Dan Wiesch (2020, p. 99), o fluxo que uma postagem segue é inicialmente desconhecido, não há uma lógica pré-definida, por isso ele argumenta que, em razão dos parâmetros de conectividade, “são extintas as possibilidades de correção existentes no interesse da sociedade como um todo.” Desse modo, demonstrar a falsidade de um conteúdo, corrigindo a informação é ineficiente para a coletividade porque é difícil refazer o caminho percorrido em rede e a revelação do falso não atinge a mesma audiência. Além disso, diante das fake news, verdade e veracidade são relegadas para o segundo plano. Isso significa que demonstrar a falsidade de uma fake news, apesar de ser relevante, não interrompe os efeitos nocivos que produzem enquanto acessíveis.

Há outra controvérsia sobre a filtragem de fake news, mesmo os doutrinadores contemporâneos divergem sobre qual a melhor concepção adotar em relação a moderação de conteúdo em virtude da possibilidade de manipulação do debate público pela autoridade responsável por estabelecer quais manifestações poderiam ou não circular com base no julgamento subjetivo de verdade ou mentira das declarações.

Para Clarisse Gross (2020), a imprecisão quantos aos critérios que definem quais os discursos são e quais não são valiosos coloca os cidadãos sob o risco de punições e proibições indevidas, especialmente nos discursos de opinião. Desse modo, para a autora, a autocensura e o efeito silenciador de vozes que se calam por temerem a responsabilização são os riscos de restringir o direito de expressão com base na verdade do conteúdo. No entanto, é importante destacar que a autora expressamente trata sobre discursos de opinião e as fake news se referem a fatos.

Adotando esse entendimento, o então Presidente Jair Bolsonaro vetou o artigo 359-O da Lei 14.197/2021 (BRASIL, 2021b) que revogou a Lei nº 7.170/193, conhecida como Lei de Segurança Nacional. No dispositivo vetado havia previsão para o crime de comunicação enganosa em massa que trazia a seguinte redação:

359-O. Promover ou financiar, pessoalmente ou por interposta pessoa, mediante uso de expediente não fornecido diretamente pelo provedor de aplicação de mensagem privada, campanha ou iniciativa para disseminar fatos que sabe inverídicos, e que sejam capazes de comprometer a higidez do processo eleitoral:

Pena - reclusão, de 1 (um) a 5 (cinco) anos, e multa.

Nas razões do veto, justificou que a proposição legislativa não foi cristalina sobre “qual conduta seria objeto da criminalização, se a conduta daquele que gerou a notícia ou daquele que compartilhou (mesmo sem intenção de massificá-la)”, se o crime seria continuado ou permanente e qual seria o órgão encarregado de definir o quê seria entendido por inverídico como um “Tribunal da Verdade”. Ainda aduziu que a redação genérica com ampla discricionariedade do intérprete em relação a conduta dolosa do verbo “promover” traria muita insegurança jurídica e afastaria o eleitor do debate político.

Para Gilmar Mendes (BRANCO E MENDES, 2020, p. 360) “O requisito da verdade deve ser compreendido como exigência de que a narrativa do que se apresenta como verdade factual seja a concussão de um processo de busca de reconstrução da realidade. Traduz-se, pois, num dever de cautela imposto ao comunicador.” Sobre essa controvérsia, Chiara Spadaccini de Teffé e Carlos Affonso Souza (2020, p. 291) defendem que:

(...)os deveres de verdade, transparência, cuidado e de checagem de fontes são dirigidos apenas a certas pessoas, como jornalistas e entidades de imprensa, em razão das normas técnicas profissionais. Em nosso entendimento, não parece que se possa afirmar a existência de um dever geral de falar a verdade para todo e qualquer cidadão.”

Discorrendo sobre a temática, Edilsom Pereira de Farias (2008, p. 146) destaca a diferença entre liberdade de expressão e liberdade de comunicação. Para o autor, a primeira é mais abrangente, liberdade de emitir opinião e juízos de valor e, portanto, não sujeita ao teste de veracidade. A segunda refere-se à responsabilidade do emissor de não difundir fatos sabidamente falsos e enganosos e ao direito de receber fatos verificados ou verificáveis.

O limite interno da veracidade, aplicado à liberdade de comunicação, refere-se à verdade subjetiva e não à verdade objetiva. Ou seja, num estado democrático de direito o que se exige do sujeito é um dever de diligência ou apreço pela verdade, no sentido de que ela seja contactada a fonte dos fatos noticiáveis e verificada a seriedade ou idoneidade da notícia antes de qualquer divulgação. (FARIAS, 2008, p. 147).

Interpretando esse trecho, pode-se extrair que, para o autor, a liberdade de opinião não é reprovável ou punível, a liberdade de comunicação, no entanto, tem intrínseca a responsabilidade com a audiência, onde um direito individual não se sobrepõe a um interesse coletivo. Então, num contexto da comunicação social, seja eletrônica ou não, é preciso cautela e enseja responsabilidade, tanto para o veículo de imprensa quanto para usuário-produtor de conteúdo, quanto ao que é veiculado, mormente quando seja fraudulento.

Apesar dos relevantes argumentos e mesmo diante da dificuldade em identificar fake news, ainda que se refira a fatos, é importante ressaltar que não se está a contestar a liberdade

de expressar nem de emitir opinião. A própria liberdade de expressão é maculada se baseada em falsidades. De igual modo, não se trata sobre a escolha de uma ideologia que deverá prevalecer ou de interpretações e opiniões sobre os fatos que devem persistir e que fazem parte do dissenso democrático.

O estudo analisa a possibilidade de exclusão das fake news disseminadas por mílias digitais no intuito de diminuir a circulação de conteúdo desinformativo assim como impedir ou minimizar os efeitos nocivos num ambiente de comunicação social com dimensões expandidas, dando mais espaço e visibilidade para notícias e informações balanceadas e produzidas segundo padrões éticos e profissionais.

Excluir fake news disseminadas por mílias digitais tem correspondência com à liberdade de informação, ao exercício da cidadania sob condições reais para evitar que cidadãos tenham suas escolhas deformadas, inclusive como medida para desestimular a atuação das mílias digitais que são remuneradas de acordo com a circulação, independentemente do teor. A filtragem algorítmica pode ser um mecanismo objetivo para remoção de fake news disseminadas por mílias digitais no sentido de restaurar a comunicação social e seus valores informativos além de impedir ou dificultar os objetivos das mílias freando a remuneração que cessa com a remoção do conteúdo.

O princípio insculpido no art. 221, I, da CF/88 que estabelece a preferência a finalidades educativas, artísticas, culturais e informativas da comunicação social, seja eletrônica ou não, é desprestigiado pelo modelo de negócios que estrutura a comunicação segundo fins lucrativos. Desse modo, em vez de prestigiar fontes jornalísticas e informações e notícias balanceadas e verificadas, opta por organizar o conteúdo de acordo com finalidades lucrativas, promovendo fake news. Por isso, a filtragem de fake news disseminadas por mílias digitais além de abrir espaço e dar visibilidade à conteúdo informativo abafado por fake news implica também diminuição de conteúdo desinformativo, favorecendo a ordem informacional e reavivando os princípios constitucionais da comunicação social.

Do mesmo modo que a liberdade de expressão é garantida como direito individual e coletivo, art. 5º IV e IX, CF/88, a liberdade à informação também é garantida no art. 5º, XIV, CF/88 e encontra fundamento no 1º, II, CF/88 relacionado ao exercício da cidadania. As informações precisam ser seguras para que a vontade popular seja formada segundo circunstâncias reais. A relevância da informação deve ser medida pela correspondência de verdade, de verificação dos fatos e não conforme a quantidade de likes ou compartilhamentos dos usuários, como ocorre no ambiente de discussão online orientado por um modelo de

negócios. Para Callejón (2022a) liberdade de expressão e liberdade de informação se relacionam da seguinte forma:

Dentre os muitos elementos de tensão que se articulam em torno da liberdade de expressão, afigura-se especialmente relevante aquele suscitado entre a liberdade de opinião e a liberdade de informação, entre a notícia e a análise da notícia, entre os fatos e seu significado, entre a realidade e a narrativa. A liberdade de expressão se assenta sobre a liberdade de informação porque está associada ao que os seres humanos conhecemos da vida, da sociedade, da política, do mundo em geral. Este conhecimento se alimenta da informação, que é a que nos proporciona uma imagem da realidade. A informação, por veraz que seja, não é a realidade, senão parte da construção social da realidade. (CALLEJÓN, 2022a, p.193)

Nesses termos, o direito à informação precede a liberdade de expressão porque se reverte em favor de toda coletividade na formação da opinião pública segundo circunstâncias reais, na formação de concessos, escolhas públicas e decisões políticas. Assim, disseminar fake news compromete o exercício regular da cidadania e fragiliza a democracia por deformar a posição do eleitor, como ocorreu nas eleições presidenciais dos EUA e na campanha do Brexit na Inglaterra (CALLEJÓN, 2022a).

Hanna Arent na obra *Verdade e Política* (1967), trata sobre a importância da verdade factual para o debate público. Ela parte do pressuposto de que os fatos são a matéria das opiniões, de modo que os fatos precedem as opiniões. Assim, segundo a autora, fatos e opiniões, apesar de serem diferentes, não se opõem, pertencem ao mesmo domínio de natureza política. Numa só frase ela resume como os fatos são relevantes para a formação da opinião e escolhas políticas, deixando evidente como se relacionam o direito à informação e liberdade de expressão, verdade factual e opinião: “A liberdade de opinião é uma farsa se a informação sobre os fatos não estiver garantida e se não forem os próprios fatos o objeto do debate.”.

Desse modo, a verdade factual se mostra imprescindível para o debate público, a partir das circunstâncias fáticas é que se desenvolvem interpretações e opiniões. A mentira não pode ter mais visibilidade do que a verdade sob o risco de desvirtuamento dos ideais cívicos e democráticos. Entretanto, Evelyn Douek (2022) pondera que também não se deve concentrar os padrões de moderação na busca de um ambiente de discurso ideal livre de fake news, de conteúdo ilícito, degradante ou violento porque é utópico e implica em irrazoável comprometimento da liberdade de expressão.

Do mesmo modo que filtragens operadas sem definição objetiva sobre as regras do que pode ou não ser excluído do ambiente online tem o condão de ocasionar restrições arbitrárias à liberdade de expressão, ainda que tenham inicialmente propósitos cívicos ou humanísticos, a circulação de fake news sem um mecanismo de contingenciamento célere também causa sérias

consequências à dignidade humana e à democracia. O âmago da pesquisa tangencia esse tormentoso equilíbrio sobre qual a medida da liberdade de expressão numa democracia conectada.

Conquanto se verifique a existência de alguns elementos em comum, tendência que parece ser irreversível, ao menos considerando a forte aproximação das tradições e culturas dos direitos europeu ocidental, anglo-saxão e latinoamericano, inexistente uma resposta universalmente válida e correta tanto sobre o conteúdo e limites em geral da liberdade de expressão, quanto sobre o modo de resolver a colisão entre a liberdade de expressão e os direitos da personalidade. Em relação a estes direitos, posição de destaque recebem a honra, a imagem e a vida privada, mas também a colisão com outros bens de hierarquia constitucional, como é o caso, por exemplo, cada vez mais emblemático, da liberdade religiosa, da proteção das crianças e adolescentes, bem como das seguranças públicas interna e externa, bastando aqui apontar para o problema da criminalidade organizada e do terrorismo. (SARLET; FILHO, 2016, p.115)

Apesar de ser difícil mesmo no plano teórico elaborar uma solução uniforme, ainda que existisse, implementá-la seria mais desafiador porque estaria sempre a depender de homens falíveis e máquinas não tão precisas, por isso nenhuma proposta é definitiva ou unânime e o estudo é arrojado porque debate, de modo acadêmico, medidas de contingenciamento de fake news, limite ao dissenso democrático e restrições à liberdade de expressão.

As fake news necessitam de avaliação de conteúdo e contexto para aferir a verdade ou não dos fatos, além disso, o propósito fraudulento de trazer vantagem ou causar prejuízo também necessita de averiguação e que muitas vezes é difícil de comprovar porque o dolo tem natureza subjetiva, imbuída no ânimo do propagador. No entanto, se a disseminação viral é promovida por milícias digitais a atividade coordenada de contas torna mais evidente a intencionalidade fraudulenta tendo em vista que a ação em conjunto para a disseminação de fake news atinge milhões de usuários em instantes e potencializa os riscos de danos irreparáveis a dignidade humana e a bens coletivos juridicamente relevantes.

Por tudo que já se escreveu, sob o ponto de vista constitucional, as fake news não se sustentam porque não são essenciais ao desenvolvimento do indivíduo, mas fruto de atividade fraudulenta, não favorecem a cidadania porque o teor é desinformativo nem são expressão de pluralismo já que alimentam guerra de narrativas e a destruição da confiança, também não produzem conhecimento, menos ainda são sinônimo de diversidade sociocultural em virtude de causarem o silenciamento de minorias. Para Gilmar Mendes (BRANCO E MENDES, 2020, p. 360) “A informação falsa não seria protegida pela Constituição por que conduziria a pseudo-operação da formação da opinião.”

A produção e a disseminação de fake news também se mostra como um exercício abusivo do direito de informar por que aquele que intencionalmente propaga não tem o direito

de desinformar e, em razão disso, as fake news não são manifestações merecedoras de guarida constitucional. Dito de outra forma, não subsiste um direito individual de desinformar que autorize as pessoas a divulgarem maliciosamente fake news.

Adriana Festugatto e Bolzan de Moraes (2021) aduzem que mesmo aqueles que divulgam fake news se valem do anonimato online e da dificuldade em provar o animus delituoso como escudo. As fake news afetam negativamente o acesso à informação, quem as recebe vê-se frustrado diante das expectativas de informações confiáveis no âmbito da comunicação social. De acordo, Gilmar Mendes (BRANCO E MENDES, 2020) é categórico ao afirmar que quem recebe notícias irreais não se vê atendido no seu direito de ser informado.

A Constituição Federal de 1988, em seu preâmbulo assegura o exercício de direitos individuais e sociais, a liberdade, a segurança e a igualdade dentre outros, numa sociedade fraterna fundada na harmonia social. O preâmbulo da Carta Magna nacional, apesar de não ter força obrigatória, indica uma percepção de convivência equilibrada dos direitos, tendo em vista que “orienta os afazeres hermenêuticos do constitucionalista” (BRANCO E MENDES, 2020, p. 108).

Nenhum direito no bojo da constituição é ilimitado, não sendo os direitos pretensões unilaterais que esvaziam a fruição por outras pessoas não parece compatível com os ditames constitucionais defender que conteúdo que se saiba falso e que tenha por objetivo trazer vantagem ou prejuízo para os outros tenha o direito de permanecer online. A interpretação da constituição deve considerar a harmonia entre as normas de modo que os direitos convivam e coexistam sem que pretensões unilaterais de uns esvaziem a fruição de direitos de outros.

A liberdade de expressão ganha relevância como direito individual com fundamento na dignidade humana, art. 1º , III, CF/88 assim como também se mostra essencial ao regime democrático e ao pluralismo político fundamentado no art. 1º , V, CF/88. Do mesmo modo que a liberdade de expressão é garantida como direito individual e coletivo, art. 5º, IV e IX, CF/88, o direito à informação também o é no art. 5º, XIV, CF/88 e encontra fundamento no exercício da cidadania, 1º, II, CF/88¹⁷, tendo em vista que as informações são essenciais para a formação

¹⁷ CF, Art. 1º A República Federativa do Brasil, formada pela união indissolúvel dos Estados e Municípios e do Distrito Federal, constitui-se em Estado Democrático de Direito e tem como fundamentos:

I - a soberania;

II - a cidadania;

III - a dignidade da pessoa humana;

IV - os valores sociais do trabalho e da livre iniciativa;

V - o pluralismo político.

da opinião pessoal e da opinião pública que condensa o pluralismo em escolhas políticas e decisões eleitorais.

Nesse sentido, nem mesmo o pluralismo, como fundamento constitucional, é ilimitado. Nem todas as ideias são admitidas mesmo numa democracia, como é o caso do Brasil. Assim, nos termos do art. 17¹⁸, caput, da CF/88, apesar de ser livre a criação, fusão, incorporação e extinção de partidos, há as ressalvas constitucionais que resguardam a soberania nacional, o regime democrático, o pluripartidarismo e os direitos fundamentais da pessoa humana. Desse modo, não é admitida a existência de partido político que atente contra essas ressalvas constitucionais, onde sequer há abertura para participação de pleitos eleitorais. Pode-se comparar esse tipo de controle com o das fake news disseminadas por milícias digitais que sequer merecem circular online, devendo ser removidas desde logo, sem que tenham espaço na esfera pública ou mesmo no ambiente de discussão online.

Stephen Holmes e Cass Sunstein (2019) argumentam que direitos são meios de auto-organização e criam deveres para o bem-estar social. Para eles (HOLMES E SUNSTEIN, 2019, p. 132), a liberdade de publicar um jornal, por exemplo, “... tem a finalidade de estimular a comunicação social e não de proteger a liberdade de expressão de indivíduos como se estivessem isolados numa ordem pré-social ou promover o isolamento ou o egocentrismo hedonista.” Tal como o pluralismo político, a liberdade de expressão também não é absoluta, mas um direito a ser harmonizado com outros.

Assim toda atividade online tem de respeitar e se compatibilizar com os fundamentos constitucionais inseridos no art. 1º do texto constitucional, especialmente em relação ao fluxo comunicacional que altera a percepção da realidade pelos cidadãos e deforma as escolhas democráticas, atingido toda a coletividade.

Diante do exposto, filtragens algorítmicas de fake news disseminadas por milícias digitais operadas mediante o controle de comportamento articulado em rede associado ao

Parágrafo único. Todo o poder emana do povo, que o exerce por meio de representantes eleitos ou diretamente, nos termos desta Constituição.

¹⁸ Art. 17. É livre a criação, fusão, incorporação e extinção de partidos políticos, resguardados a soberania nacional, o regime democrático, o pluripartidarismo, os direitos fundamentais da pessoa humana e observados os seguintes preceitos:

I - caráter nacional;

II - proibição de recebimento de recursos financeiros de entidade ou governo estrangeiros ou de subordinação a estes;

III - prestação de contas à Justiça Eleitoral;

IV - funcionamento parlamentar de acordo com a lei. (...)

controle de conteúdo e contexto das notícias e informações estão dentro das possibilidades editoriais dos curadores digitais que têm por dever ser confiáveis, contudo, as filtragens devem ser imparciais e não devem ser operadas segundo vieses discriminatórios ou ideológicos por violarem artificialmente o ambiente democrático e desequilibrarem as forças que disputam o poder.

Para Evelyn Douek (2022), mesmo a moderação de conteúdo online deve ser baseada em legalidade, responsabilidade processual e transparência e não deveria se concentrar em análises ex post mas se operar de modo sistêmico, no entanto para atender as exigências de legalidade devem ter uma estrutura organizacional interna para fazer cumprir suas regras. Ela sugere uma separação do pessoal de acordo com as diferentes funções e incentivos de modo que os moderadores sejam separados dos lobistas e daqueles que se ocupam do crescimento do serviço ou da receita de publicidade.

No tocante a responsabilidade processual, a autora (DOUEK, 2022) sugere que a moderação observe mandatos estruturais imparciais para fazer cumprir suas regras, com a demonstração de sistemas capazes de implementar e supervisionar os próprios critérios de moderação, lato sensu, tendo em vista a desconfiança generalizada de que essas empresas não aplicarão suas regras da forma como declaram publicamente. Além disso, que seja feita a identificação dos moderadores para que possam ser responsabilizados por suas decisões, no intuito de limitar a discricionariedade privada.

Com relação a transparência, a autora sugere (DOUEK, 2022) relatórios mais detalhados não apenas quanto ao número de processos analisados ou a quantidade de conteúdo removido, mas também considerando como as métricas foram estabelecidas, em quais lugares a incidência do discurso infrator foi maior, por exemplo, e um plano de avaliação de risco ex ante sobre os impactos da moderação e modulação de conteúdo no intuito de trazer à tona decisões sobre o desing dos sistemas. Além disso, um canal aberto para os usuários também participarem da moderação alertando sobre falhas contínuas.

Ao avaliar a governança da internet sobre três princípios processuais do Estado do Direito, aproximando-a do constitucionalismo digital, Nicolas Suzor (2018) esclarece que as expectativas de legitimidade devem estar presentes no ambiente privado das redes mas sem o rigor que são exigidos aos entes governamentais. Desse modo, consentimento significativo, legalidade formal e devido processo são preocupações centrais na governança na internet, especialmente quanto a previsibilidade normativa e a segurança para os usuários, no intuito de limitar os que estão em posição de poder a não agir de maneira arbitrária.

Diante disso, Suzor (2018) afirma que a legalidade formal é o valor basilar do Estado de Direito e implica que o poder seja exercido de maneira responsável, que os que assumem posições de poder cumpram as próprias regras e que só haja alteração das mesmas por procedimentos apropriados.

No tocante a legalidade, o ponto central para o autor se refere a aplicabilidade e interpretação das normas de modo não discriminatório ou tendencioso. Para Suzor (2018), os termos de serviços não são claros e têm previsão de alteração unilateral, alguns com cláusula prevendo o direito a remover conteúdo por critério exclusivo a qualquer tempo com ou sem aviso prévio. O autor reconhece o desafio de estabelecer regras claras num ambiente dinâmico de modo que os usuários estejam sempre informados sobre as mudanças operadas pelas redes no intuito de se adaptarem à dinâmica online disruptiva.

Para Suzor (2018), o consentimento significativo é especialmente delicado na governança privada da internet porque os contratos normalmente estipulam muitos direitos para as empresas, mas quase nenhum dever, escapando ao máximo de responsabilidades, subjugando usuários quanto ao acesso e continuidade da prestação de serviços de maneira unilateral. Ele afirma que as cláusulas contratuais são muitas vezes incompreensíveis por serem técnicas.

Em relação ao devido processo, o autor (SUZOR, 2018) menciona a necessidade de um mecanismo para resolver disputas, com via de recurso e revisão independente, além de transparência quanto aos critérios usados pelos moderadores de conteúdo, no entanto, ele observa em seu estudo que, nos termos de serviços, prevalece o sigilo sobre os critérios de moderação de conteúdo havendo poucas chances de um processo formal interno.

Assim, as filtragens algorítmicas para excluir fake news disseminadas por milícias digitais apesar de embasadas nos princípios constitucionais da comunicação social, especialmente, em relação a preferência a atividades informativas devem ser imparciais, sem atuação enviesada para que as remoções de fake news de uma específica vertente acabe por favorecê-la, ou seja, evitando salvaguardar a verdade dos fatos somente para um candidato ou ideologia ao tempo em que negligencia ou até mesmo estimula fake news que abalem a credibilidade dos adversários.

A comunicação social sempre esteve em mãos privadas que realizam controles internos não havendo óbice para que as empresas gestoras dos fluxos comunicacionais também procedam contingenciamentos em benefício da coletividade, ainda que a moderação recaia sobre conteúdo de terceiros porque o contingenciamento de fake news disseminadas por

milícias digitais não se enquadra dentro das garantias da liberdade de expressão individual nem dos valores democráticos.

A questão então não seria se as exclusões de fake news disseminadas por milícias podem ou não ser operadas proativamente pelas empresas, mas se a forma como se dá esse controle é imparcial, transparente e confiável. Desse modo, para garantir as regras do jogo democrático do controle online de fake news disseminadas por milícias digitais, entendo que seriam observadas regras procedimentais relacionadas a transparência, manutenção do dissenso no mesmo equilíbrio da realidade física e possibilidade de recorrer da decisão, por isso, é tão importante regular formalmente esse critério de moderação mediante lei, para que não seja uma voluntariedade ou serviço adicional, mas uma obrigatoriedade e dever das empresas gestoras dos fluxos comunicacionais perante a sociedade.

A partir do entendimento que as empresas gestoras dos fluxos comunicacionais são curadores digitais e responsáveis pela estruturação e organização das informações e notícias, naturalmente lhes são atribuídas obrigações de supervisão porque o poder de conduzir os fluxos comunicacionais sociais implica responsabilidades não somente diante dos usuários-contratantes, mas perante a coletividade e seus interesses jurídicos relevantes.

A atuação de milícias digitais usando a infraestrutura da internet para pulverizar fake news promovendo teorias da conspiração, discursos extremistas e disseminando brutalidade nas comunicações não tem condão de promover nem a dignidade, nem de informar ou refletir pluralismo democrático, nem disseminar conhecimento ou interação entre os cidadãos. Balkin (2020) é categórico ao afirmar que o objetivo das fake news é destruir a confiança.

Ao promover desinformação, semeando dúvida e desarticulando as redes de cooperação, promovendo a distorção da realidade compartilhada em escala, segundo propósito de causar prejuízo ou aferir vantagens, as milícias digitais atentam contra elementos de coesão social, porque atuam sem qualquer comprometimento com a convivência social, com a política ou com a verdade factual e assim radicalizam e fomentam guerras de narrativas que deturpam a compreensão social sobre a realidade, maculam as decisões públicas ou mesmo impedem a construção de decisões compartilhadas.

A exclusão de fake news disseminadas por milícias digitais pode ser operada pelas empresas de tecnologia tanto porque são atividade fraudulenta ou porque não servem ao desenvolvimento da pessoa e portanto não contribuem para a dignidade da pessoa humana, como também porque atentam contra os valores da liberdade de expressão e os princípios da

comunicação social. Há fundamentação jurídica suficiente para embasar as remoções se praticadas desvinculadas de interesses lucrativos sem vieses ou tendências.

Repetidas vezes, o estudo se referiu às práticas das empresas gestoras dos fluxos comunicacionais como obscuras, sombrias e não confiáveis e, por isso, foi reforçada a necessidade de lei para instituir a exclusão de fake news disseminadas por milícias digitais como critério obrigatório de controle de conteúdo online, prescrito em lei, observadas a imparcialidade, o equilíbrio do dissenso no ambiente democrático tomando como referência a realidade física e o direito a impugnação com resposta fundamentada e célere. Ocorre que, em busca de uma paridade de armas, têm circulado ideias extremadas como a democracia defensiva, que consiste numa forma radical de retirar do cenário democrático elementos subversivos.

O Brasil, segundo Tarsila Fernandes (2021), adotou a teoria da democracia de militância no art. 17 da CF/88, então, mesmo o país tendo como fundamento o pluralismo, não admite a criação de partido político que atente contra o regime democrático. Nesse sentido, a teoria entende que a democracia pode repelir uma ameaça real que se vale dos mecanismos democráticos para corroê-la ou destruí-la. Assim, de acordo com a autora, a democracia de militância é um controle interno para frear ideologias e organizações, como os partidos políticos, de participar do cenário político ou de um processo eleitoral:

As premissas da democracia militante, portanto, eram a de que o regime democrático deveria contar com mecanismos (ainda que antidemocráticos) para evitar que agentes políticos com ideais totalitários de poder, tais como Hitler, utilizassem instrumentos democráticos para chegar ao poder. Assim, deveriam ser criados meios para que a democracia se defendesse dos partidos que buscassem alçar-se ao poder para destruí-la. (FERNANDES, 2021, p.135)

Diferentemente da democracia de militância, a democracia defensiva não foi prevista constitucionalmente no país. Desse modo, ampliar margens de interpretação judicial no tocante a avaliação de abusos de direito de modo a conferir poderes ao Judiciário Brasileiro para suprimir esses mesmos direitos sem que lei defina quais as hipóteses são possíveis esse controle implica criar uma fórmula temerária e antidemocrática que concentra intensa carga subjetiva numa autoridade.

Seguindo a mesma racionalidade da democracia de militância, a teoria da democracia de resistência também foi desenvolvida para impedir que grupos extremistas, situações indesejadas ou momentos de instabilidade institucional afetem o regular funcionamento da democracia. Apesar de terem a mesma origem, a teoria da democracia defensiva diferencia-se da democracia de militância por ser mais ampla. Vale dizer, a democracia de militância foi criada e desenvolvida para lidar com uma situação específica: a questão da exclusão de partidos ou de grupos políticos totalitários do processo democrático. A democracia defensiva, entretanto, visa

impedir que vulnerem a própria democracia quaisquer situações ou grupos que afetem a normalidade democrática. (FERNANDES, 2021, p.138)

Apesar de ter como fundamento constitucional o pluralismo, a proibição de criação de partido político que atente contra o regime democrático, nos termos do artigo 17 da CF/88 denota que nem toda dissidência é permitida, que há um limite. A teoria da democracia de militância em 1937, como medida não democrática, foi desenvolvida na Alemanha. Passados mais de 50 anos foi adotada no Brasil.

O que antes não era convencional foi sedimentado na CF/88 anos mais tarde em razão dos perigos das ideologias totalitárias. Do mesmo modo, a defesa da filtragem algorítmica de fake news disseminadas por milícias digitais pode parecer não ortodoxa se analisada de modo superficial mas é medida assecuratória frente a ataques engendrados que acionam algoritmos tendenciosos ao lucro numa governança autoritário e essas práticas fragilizam a democracia e abrem espaço para regimes despóticos a medida que destroem a confiança nas instituições e deformam a leitura da realidade socialmente compartilhada.

Como as fake news disseminadas por milícias digitais não se compatibilizam com pluralismo político, liberdade de expressão e de informação, a filtragem algorítmica deve ser considerada como medida necessária tendo em vista que, nessa circunstância, o limite do dissenso é excedido se considerados como parâmetros da avaliação os perigos das práticas online, o cenário de involução democrática, e até mesmo de incomunicação social eletrônica, além das interferência em processos eleitorais como ocorreu nos EUA.

Ocorre que a situação em análise na dissertação é a exclusão de fake news disparadas por milícias digitais feita pelas próprias empresas gestoras dos fluxos comunicacionais online dentro de seus respectivos domínios e, apesar de exercerem a função de regular o discurso no âmbito da comunicação social online, são pessoas jurídicas privadas que atuam mediante a cláusula geral de legalidade, diferentemente dos poderes constituídos que devem observância a controles constitucionais no modo de atuação.

Desse modo, ampliar efeitos de decisão colegiada sobre desinformação com margem para interpretação judicial no tocante a avaliação de identidade de conteúdo de fake news de modo a conferir poderes ao Presidente do Tribunal Superior Eleitoral para suprimir direitos constitucionalmente, como devido processo legal, ampla defesa, contraditório, juiz natural dentre outros, assegurados em face o poder público implica criar uma fórmula temerária e antidemocrática que concentra intensa carga subjetiva numa só autoridade judiciária.

Diante das transformações de um mundo globalizado e interconectado, Callejón (2022b) entende que não haverá mais um sistema jurídico unitário baseado nos princípios da unidade, coerência e plenitude. Para ele, a realidade está dividida, entre a física e a virtual, de modo que a realidade virtual submetida ao direito privado, não tem nada a ver com o sistema de fontes estatais. Assim, não será mais possível reconstituir a unidade porque conviverão regras de direito privado a nível global e normas de direito público a nível estatal.

Do mesmo modo, o autor (CALLEJÓN, 2022b) entende que a coerência não poderá ser realizada plenamente porque haverá conflitos a nível global que estarão entre as realidades física e virtual e serão resolvidos por cláusulas contratuais fora do sistema jurídico estatal tendo em vista a ascensão das grandes empresas de comunicação global como lideranças globais.

Uma nova arquitetura mundial está se desenhando, com a imposição de práticas concretas que não se pode ignorar, como a interferência cada vez maior das empresas gestoras dos fluxos comunicacionais online tanto no controle de conteúdo quanto no comportamento de milhões de usuários. Para enfrentar as crises dos modelos atuais de produção do direito e de solução dos conflitos pelo Estado, o primeiro passo é entender como as dinâmicas estão se materializando para conceber soluções, ainda que não procedimentais, analisando os fundamentos, valores e direitos constitucionais no intuito de planejar as mudanças e teorizar estratégias de defesa da democracia porque apesar de ser um regime resiliente precisa ser protegido.

Desse modo, a filtragem de fake news disparadas por milícias digitais apesar de não estar revestida num conceito firme do que sejam milícias digitais ou fake news bem como serem operadas sem um procedimento que previamente define suas etapas, se justifica pelo resultado de não agravar o quadro de involução democrática.

A teoria da democracia de militância se justificou no contexto brasileiro e foi expressamente incorporada na CF/88, mesmo sendo medida procedimentalmente não democrática, porque visa proteger a democracia ao impedir a participação e até a criação do partido político nos termos do art. 17. No entanto, aplicação da teoria da democracia defensiva ou de resistência, ainda que esteja condicionada a atuação mediante um dos Poderes constituídos, sem previsão constitucional expressa, é precária, legalmente frágil e perigosa porque concede poderes ao Presidente do TSE, mediante resolução do próprio Tribunal, no sentido de estender os efeitos de decisão plenária, com identidade de conteúdo, de modo abrangente e com prazos fulminantes. Além disso, não favorece a democracia sendo medida

extremada e temerária, inclusive porque protege interesses da própria autoridade judiciária e faz diferença no tratamento dos controles de fluxos em relação aos cidadãos e candidatos.

6 CONCLUSÃO

Ao longo da dissertação foi analisado que a exclusão de fake news disseminadas por milícias digitais não são limitações arbitrárias à liberdade de expressão, porque as fake news não contribuem para a formação da personalidade nem para um debate cívico plural, na medida que os discursos online envolvidos por fake news acabam por prejudicar as condições de diálogo democrático. Discussões baseadas em fake news não progridem em decisões conciliadas, porque as interações ficam condicionadas a um contínuo desvendar do falso diminuindo as chances de trocas argumentativas ou de reflexões.

As fake news disseminadas por milícias digitais têm o potencial ainda mais destruidor para o debate público porque as pessoas são expostas a ataques sistemáticos de grupos articulados em rede. O comportamento dos perpetradores usando contas inautênticas e disseminadores artificiais dá maior visibilidade ao conteúdo semeando dúvida e causando instabilidade social ao atingir milhares de pessoas em pouco tempo. Portanto, a exclusão algorítmica de fake news de disseminadas por milícias digitais não configura restrição indevida à liberdade de expressão.

Ao abordar fake news tem-se que estas podem se referir a fatos falsos ou podem ser informações verdadeiras incompletas, descontextualizadas, relativas a fatos pretéritos, mas apresentadas como algo novo/atual, ou ainda enviesadas, por isso a comprovação do falso é laboriosa e cara, tendo em vista que os aperfeiçoamentos tecnológicos aprimoram as técnicas de produção e foram absorvidos pelas atividades fraudulentas. Além disso, a verdade factual foi relegada diante das convicções pessoais, por isso, apenas desconstruir ou descortinar as narrativas das fake news não é mais suficiente.

Diante de uma comunicação em massa conectada e supersônica, a preocupação está mais atenta aos efeitos arrebatadores sobre a audiência, especialmente tendo em vista que os mecanismos constitucionais de reparação como direito de resposta e indenização dos danos são insuficientes frente às dimensões virtuais, à proporção dos danos, à dificuldade em identificar o perpetrador em razão do anonimato online e a condição de hipossuficiência econômica, além da dificuldade em refazer os caminhos que a fake news percorreu numa rede interligada por infinitos nós.

A liberdade de expressão individual é essencial para a formação de uma opinião pública plural, no entanto, a disseminação de fake news por milícias digitais não configura exercício da liberdade de expressão, mas a instrumentação desse direito por criminosos e desse modo não contribui para um debate democrático porque a intenção é enganar para auferir vantagem ou causar prejuízo. Ocorre que é difícil a comprovação do caráter fraudulento porque faz parte do ânimo subjetivo de quem divulga, mas o comportamento articulado de vários perfis sincronizados para disseminação de fake news no intuito de atingir milhares de pessoas em pouco tempo torna muito mais evidente a natureza delituosa.

O trabalho primeiro analisou a comunicação online e as nuances que a distingue de uma interação ao livre, diferenciando o ambiente de discussão online de esfera pública habermasiana, distinguindo a participação online de uma expressão cívica na democracia direta, tendo em vista a ausência de trâmites ou a impossibilidade de os cidadãos participarem e se informarem sobre todos os assuntos relacionados à sociedade.

A dissertação também apontou o contraste da internet atual que, apesar de ser um ambiente plural prioriza fake news porque as empresas que fazem a gestão dos fluxos comunicacionais online privilegiam conteúdo que promove a agitação e gera engajamento como compartilhamentos, comentários e curtidas porque mais lucrativo. O modelo de negócios consiste em captar a atenção dos usuários por mais tempo, viabilizando a coleta de mais dados para serem usados para modulação de conteúdo, direcionamento de publicidade e mesmo para influenciar ou persuadir comportamentos.

Além disso, caracterizou as análises algorítmicas como efetivamente normas de comando geral, demonstrando e comprovando que não há uniformidade entre os critérios de moderação de conteúdo entre as empresas ou mesmo entre as redes, depois investigou como a liberdade de expressão e as informações têm sido manejadas segundo um projeto de instrumentalização da humanidade, apontando a necessidade de proteção sistêmica da audiência submetida a ataques de fake news por milícias digitais que constroem artificialmente falsos cenários.

O estudo diferenciou também o modelo de governança da internet do regime democrático considerando as características procedimentais de Bobbio e caracterizou as fake news, abordando contexto e conteúdo, trazendo aspectos como conceito, finalidades, motivos pelos quais prosperam e os efeitos, bem como analisou como o controle de conteúdo caso a caso é ineficiente diante do volume de manifestações diárias, das possibilidades tecnológicas e da necessidade de medidas céleres de contingenciamento.

De igual modo, descreveu a atividade miliciana e o potencial ainda mais lesivo do comportamento articulado em rede porque, mais do que a mera divulgação de fake news, as milícias têm, muitas vezes, como finalidade específica criar artificialmente falsos cenários. Também explicou como a filtragem algorítmica é um mecanismo eficiente de controle e, apesar de apresentar falhas, ainda pode ser usada para impedir ou minimizar os efeitos danosos das fake news, viabilizar mais espaço para notícias e informações verificadas, balanceadas e produzidas segundo padrões profissionais, além de interromper a circulação e a respectiva remuneração desestimulando esses comportamentos milicianos.

As fake news disseminadas por milícias digitais tem o potencial destruidor da democracia e essa visão de perigo não é baseada num ou noutro episódio, mas na sistemática dos ataques de magnitudes mundiais que se aprimoram cada dia mais, com técnicas de manipulação de comportamento sublimar, articulação internacional e um propósito direcionado, não meramente um interesse financeiro ou uma vantagem. Portanto, o contingenciamento de fake news com exclusões de conteúdo é uma medida necessária mesmo que não convencional. Ainda que a filtragem algorítmica seja operada fora dos poderes constitucionais é um controle com potencial de eficiência e efetividade tendo em vista serem as próprias empresas as únicas detentoras de códigos operacionais aptos para excluir de modo abrangente as fake news online e estancar ou minimizar seus efeitos embora não seja tão preciso e admita falhas.

A exclusão de fake news disseminadas por milicias digitais é dinâmica que associa o conteúdo e contexto das mensagens ao comportamento dos perfis em rede para proceder as filtragens. Mesmo a demora de minutos para que se operem os filtros pode ser tardiamente comprometedor tendo em vista a concretização dos efeitos e dos danos, por isso ainda que seja um julgamento preliminar se mostra compatível com a aceleração social que se tornou marca contemporânea tendo em vista que os usuários respondem de modo reativo e não reflexivo.

Além disso, restringir conteúdo não é medida inédita na comunicação social, a mídia tradicional como rádio e televisão já fazia escolha editorial ao selecionar pautas descartando temas e abordagens problemáticas mediante decisões unilaterais do editor-chefe. Então, sob o aspecto da comunicação social não é novo o controle por pessoas privadas, inéditas são as dimensões e as ferramentas tecnologias de filtragem de conteúdo.

Ocorre que na mídia tradicional como rádio e televisão, as restrições incidiam sobre o conteúdo produzido pela própria organização jornalística e por isso não se questionava a arbitrariedade. A diferença é que, na internet, as remoções e bloqueios se operam sobre

manifestações dos usuários por isso longo da dissertação a liberdade de expressão foi analisada segundo os valores democráticos e conclui-se que as fake news não estão protegidas pelas garantias da liberdade de expressão, além de serem contrárias à liberdade de informar, como direito de informar ou de ser informado.

As práticas operacionais online superdimensionam fake news, fazendo com que teorias de conspiração, ataques aos direitos da personalidade, afrontas infundadas às políticas de saúde pública e insultos a própria ordem democrática dentre outros temas tenham uma repercussão aumentada por uma lupa. Assim, minorias e extremistas têm uma repercussão ampliada se comparada a realidade física, trazendo a falsa sensação de que se vive uma atmosfera de iminente golpe, reviravolta e instabilidade.

Apesar de não ter encontrado um limite do dissenso, sabe-se que existe um nível para tensões que a democracia suporta e nesses termos as fake news disseminadas por milícias digitais foram consideradas elementos subversivos às regras do jogo democrático por estimularem a conflituosidade e polarização em escala ao invés da consenso e do entendimento. Além disso, desequilibram artificialmente às forças que disputam o poder porque não há argumento capaz de as contrapor, em igualdade de condições tendo em vista que o comportamento humano racional e reflexivo não é suficiente para frear fake news online disseminadas por milícias digitais porque a atividade coordenada em rede dispara os mecanismos algorítmicos de impulsionamento, além disso o uso de robôs incrementa a propagação de maneira sobre-humana. Depois porque as máquinas, perfis inautênticos, disseminadores artificiais e perpetradores não internalizam reflexões.

Assim, se a percepção sobre os efeitos fake news é caso a caso, a resiliência da democracia não parece estar em risco, no entanto, diante de uma visão sistêmica o perigo à democracia é realçado considerando a quantidade e a sistemática reiterada das fake news disseminada por milícias digitais, articuladas inclusive internacionalmente.

Então, apesar de não se conhecer qual o limite do dissenso que a pluralidade é capaz de absorver e processar em instâncias democráticas, nem qual o real efeito que as fake news implicam para a democracia é plausível argumentar que fake news disseminadas por milícias digitais escapam desse limite porque a natureza fraudulenta é acentuada e ainda potencializada por comportamentos em rede articulados associados a mecanismos tecnológicos como contas inautênticas e disseminadores artificiais que agravam o cenário de involução democrática, de incomunicação social e de interferência em processos eleitorais.

Além da atividade miliciana organizar uma ação coordenada, a lógica algorítmica das redes, que persegue lucros, atua favoravelmente às fake news, impulsionando conteúdos que atingem emoções e causam agitação online como compartilhamentos, comentários e curtidas. Diante do modelo de negócios que faz a gestão dos fluxos comunicacionais online já se constata um quadro de involução democrática, retrocesso na comunicação social e forte influência das fake news nos processos eleitorais.

Apesar da remoção das fake news online serem operados por empresas privadas com base num julgamento rápido, considera dois fatores, tanto o conteúdo quanto o comportamento e, portanto, abrange duplo critério para moderação reduzindo as chances de arbitrariedades. Ocorre que o controle das fake news por filtragens algorítmicas não segue processos e ritos prévios estabelecidos por instâncias políticas e esse é um dos dilemas. A filtragem sistêmica admite falhas e possibilidades de erros como restrição indevida da liberdade de expressão. Essas encruzilhadas fazem parte do cenário do mundo conectado.

A democracia é essencialmente procedimental com trâmites definidos trazendo uma previsibilidade dos cursos. O controle online dos fluxos online não segue uma cadência linear, nem tem os critérios rigidamente definidos, é difuso e aberto, próprio para o ambiente em rede. A flexibilidade, a adaptabilidade, a absorção da inovação de forma instantânea e mesmo a disrupção fazem parte do dinamismo em rede então a margem de discricionariedade no ambiente virtual tende a ser dilatada para que a moderação esteja em compasso às práticas online e para que as filtragens operadas por inteligência artificial possam se desenvolver na mesma medida que as atividades criminosas online.

Assim, a internet possibilita inúmeras vantagens como acesso ao conhecimento, pluralidade de fontes, maior participação e liberdade, mas também tem ocasionado a involução democrática com seu modelo de negócios, no entanto, o processo de digitalização da sociedade é irreversível e, diante do volume de dados online, o processamento célere das informações e notícias depende de tecnologia. Ainda que a internet e o modelo de negócios online tenham trazido efeitos deletérios como as fake news não é concebível viver num mundo sem internet ou sem gestores dos fluxos comunicacionais para organizar e estruturar as informações e notícias.

Desse modo, apesar das falhas e dos erros dos controles de conteúdo sistêmicos também se tornou desarrazoado defender que o modelo padrão de moderação de conteúdo permaneça ex post e caso a caso diante da quantidade de dados e do potencial nocivo das fake news tendo em vista o número de pessoas atingidas. Como os efeitos em rede são exponencialmente

ampliados então os danos, os prejuízos e as consequências se tornam ainda mais gravosos pelas dimensões virtuais. As fake news disseminadas por milícias digitais tem um alcance expandido e as filtragens algorítmicas são aptas para cessar a propagação e mesmo excluir conteúdo a fim de evitar os efeitos como os do genocídio como o de Mianmar e o da influência negativa em campanhas de imunização como no caso da COVID ou interferência em eleições como ocorreu nos EUA, por isso são importantes ferramentas para a manutenção da própria ordem informacional e da democracia.

Assim, a análise de constitucionalidade considera uma leitura dos dispositivos constitucionais em consonância com o quadro atual, com os perigos que as fake news disseminadas por milícias digitais representam e com os riscos da estrutura comunicativa da internet à estabilidade social e à democracia, tendo em conta que o Estado não tem capacidade de proceder controle de conteúdo online.

Esse cenário de moderação de conteúdo online já se incorporou ao cotidiano e se concretiza nas diversas práticas online pelas empresas que fazem a gestão dos fluxos comunicacionais com diferentes normas internas, cada qual com seu sistema de cláusulas contratuais, termos de serviços e diretrizes da comunidade que perfazem diversos regramentos próprios, como demonstrado.

No Brasil, as limitações constitucionais sobre a liberdade de expressão foram desenhadas para um modelo de comunicação social de massa tradicional como rádio e televisão e se mostram ineficientes e obsoletas diante das novas tecnologias. São tempos de ruptura, real e virtual são mundos diferentes, com possibilidades diferentes e nitidamente há um avanço do mundo virtual sobre espaços reais, comprimindo a realidade física, de modo que afeta não apenas direitos, mas a política e a economia, inaugura novos arranjos sociais e uma outra distribuição de poder não mais centralizada no Estado mas partilhada com sistemas supranacionais e com atores privados como as Big Tech.

As milícias digitais usam a infraestrutura em rede e os mecanismos algorítmicos de ampliação do alcance de um conteúdo para promover ataques coordenados, assim ao disseminar fake news de forma maciça e pulverizada destroem a realidade socialmente compartilhada semeando desconfiança e dúvida com força para influenciar ou mesmo determinar comportamentos usando artifícios subliminares para despertar emoções como raiva, medo e repulsa.

Ao disseminar fake news, semeando dúvida e desarticulando as redes de cooperação sociais, as milícias digitais atuam com o propósito de causar prejuízo ou aferir ganhos financeiros sem qualquer comprometimento com a convivência social ou com a verdade factual, radicalizando opiniões e fomentando guerras de narrativas que deturpam a compreensão social sobre a realidade, maculando as decisões públicas ou mesmo impedindo a construção de decisões compartilhadas, ficando o debate público preso num ciclo infundável de descortinamento e elucidação dos fatos.

Para os gestores da comunicação online, que são grandes empresas privadas, não há proibição para a prática da moderação de conteúdo que conjuga os dois parâmetros, comportamento e conteúdo, no entanto, é preciso que estejam definidas as hipóteses de incidência e que as filtragens sejam imparciais e sem vieses discriminatórios. Ainda que não exista uma fonte central de regulação tendo em vista que cada empresa traz um conceito diferente para milícias, com padrões nem sempre coincidentes, moderando com maior ou menor grau essa atividade coordenada online, o controle sobre a atividade miliciana de disseminar fake news se mostra mais compatível com a dinâmica online do que o modelo ex post, caso a caso, como o judicial.

Ocorre que, em virtude do volume de dados que circula online, a gestão dos fluxos comunicacionais se opera em grande parte através de análises algorítmicas razão pela qual parece contraditório o poder público esperar que, na moderação de conteúdo online feita pelas empresas, as garantias processuais que vigoram em relação aos cidadãos sejam estendidas, na mesma proporção, aos usuários. Isso seria impraticável tendo em vista a escala de manifestações e quantidade de moderações diárias.

Desse modo, as garantias processuais constitucionais não são transpostas do âmbito público para o privado, entre usuários e empresas, com o mesmo rigor. A moderação de conteúdo do online confere uma maior discricionariedade aos moderadores humanos e aos programadores para que a regulação possa se atualizar diante das dinâmicas digitais que são disruptivas e para seja compatível à flexibilidade do ambiente online. Assim, a legalidade estrita é mitigada e preenchida por normas técnicas produzidas nos domínios privados das próprias redes. O devido processo legal também se adequa ao devido processo curacional (curadores digitais), sendo assegurado o direito de resposta à reclamação e em prazo razoável, mas sem um sistema recursal complexo dividido em instâncias.

Desse modo, não haveria pelo controle sistêmico automatizado de filtragens de fake news disseminadas por milícias digitais a exposição de razões individualizadas ou

comunicações prévias sobre remoção do conteúdo. Assim, para garantir as regras do jogo democrático devem ser observadas regras procedimentais relacionadas a transparência, manutenção do dissenso no mesmo equilíbrio da realidade física e possibilidade de recorrer da decisão, por isso, é tão importante regular formalmente esse critério de moderação mediante lei, para que não seja uma voluntariedade ou serviço adicional, mas uma obrigatoriedade e dever das empresas gestoras dos fluxos comunicacionais perante a sociedade.

As filtragens algorítmicas para exclusão de fake news disseminadas por milícias digitais têm de ser imparciais, ideologicamente desinteressadas, sob pena de causar desequilíbrios democráticos e, ao invés de servirem como meios de correção e restauro do ambiente de debate, as filtragens serviriam a usuários economicamente privilegiados que poderiam pagar por esses serviços em rede no intuito de remover conteúdo que lhes seja desfavorável. Portanto, as filtragens não devem ser um serviço ou liberalidade a ser oferecida aos usuários ou meramente um direito das empresas gestoras dos fluxos comunicacionais, mas dever diante dos riscos e responsabilidade pelos danos provocados à terceiros decorrentes suas atividades lucrativas.

Havendo alteração da situação de incapacidade técnica do Estado, para proceder moderação de conteúdo simultânea aos fluxos, as conclusões do estudo podem ser revistas, seja porque houve o investimento e desenvolvimento de tecnologia para diretamente intervir nos fluxos online ou mesmo porque o treinamento de corpo administrativo para fiscalizar e atuar no interior das empresas gestoras da comunicação global capacita às autoridades públicas e lhes atribui poder para orientar o desenvolvimento de sistemas e filtragens segundo propósitos e finalidade públicas.

Enquanto persistir a realidade de que os tribunais e seus ritos não são efetivos nos controles de conteúdo online, especialmente porque muitas vezes sequer conseguem identificar os perpetradores, distribuídos mundo afora, como é comum em atividades milicianas, a moderação de conteúdo feita pelas próprias empresas é importante ferramenta para proteger a democracia. As formas de combate analógicas como direito de resposta e indenização pelo dano sofrido mediante apuração caso a caso pelo Poder Judiciário, após um trâmite processual minucioso, não se compatibilizam com as dinâmicas online tendo em vista a extensão e a permanência dos efeitos.

Ainda que tenha sido concedido o direito de resposta ou realizado o pagamento da indenização, as fake news continuam a produzir danos enquanto permanecem online, podendo ressurgir a qualquer tempo, ressuscitando os mesmos fatos falsos só que divulgados por outros

perfis ou contextualizados para atingir outras pessoas, o quê demanda novos processos judiciais ou pedidos para extensão dos efeitos de decisão colegiada numa cadeia sem fim.

As altas velocidades de transmissão digital e o ritmo de compartilhamento instantâneo, como cliques, impõem celeridade às medidas de contingenciamento, de modo que sejam capazes de reagir na mesma proporção para equilibrar a atividade miliciana e os mecanismos de controle. Como o Estado não tem capacidade para proceder a moderação de conteúdo online, essa função acaba sendo exercida pelos gestores dos fluxos comunicacionais, empresas privadas. No entanto, instrumentos como a Resolução 23.714/2022 do TSE são temerários ainda que sob a justificativa de estender os efeitos de decisão colegiada para atender na segurança jurídica, coerência, efetividade e agilidade às decisões judiciais.

Se a moderação de conteúdo operada em espaços privados requer lei para garantir a imparcialidade e o aperfeiçoamento para que os resultados sejam cada mais precisos de modo a evitar filtragens com vieses, tendências, voluntarismos ou para que não seja oferecida como serviço a disposição de quem puder pagar, beneficiando os economicamente privilegiados. Com muito mais razão, deve haver embasamento legal para negativa e supressão, pelo Poder Público, de direitos constitucionalmente previstos segundo a teoria da democracia defensiva.

Assim, apesar da disseminação viral por milícias digitais promover agitação online ao atingir muitos perfis em poucos instantes e ser, a princípio, financeiramente interessante, a permanência das fake news torna o ambiente carregado de toxidade e causa a evasão dos usuários e anunciantes o quê faz despencar o potencial lucrativo do negócio, portanto, as empresas também têm interesse numa moderação de conteúdo equilibrada, mantendo o ambiente plural e sem a perspectiva de uma higidez absoluta de seus domínios porque isso a tornariam menos atrativas para a comunidade dos usuários.

As alegações de que remoções de conteúdo ferem a liberdade de expressão são superficiais e generalizam todas as manifestações como legítimas. Ocorre que fake news não se encaixam nos valores da democracia política, cultural ou de disseminação de conhecimento, assim como as fake news não atendem o direito de informação nas duas dimensões, direito de informar e de ser informado, tendo em vista que nem todos tem acesso as mesmas informações em razão das modulações e personalização do feed de notícias.

Do mesmo modo, fake news também não são expressão de pluralismo político porque a diversidade da realidade física não tem a mesma proporção no ambiente online em virtude de serem amplificadas por análises algorítmicas e as minorias serem representadas por meio de

lentes de aumento, fazendo crer que são majoritárias. Então, excluir fake news disseminadas por milícias digitais online não significa suprimir vozes divergentes e dissonantes que continuarão a existir independentemente das filtragens.

As fake news têm natureza fraudulenta e por essa razão o tratamento jurídico não deve ser o mesmo que o conferido a manifestações cujo propósito é expressar e contribuir para o desenvolvimento humano, participar das decisões políticas, contestar, denunciar ou reivindicar atividade estatal.

As gestoras da comunicação online assumem os riscos de danos ao modular e moderar conteúdo de acordo com interesses lucrativos, dando mais visibilidade às fake news razão pela qual a tarefa de monitorar seus espaços passa a ser, além de um direito já que não há proibição expressa, uma obrigação tanto para proteger a democracia quanto para evitar o esvaziamento de sua comunidade.

Desse modo, modular e moderar conteúdo também ocorre no interesse do modelo de negócios, no entanto, a composição entre o interesse privado e o público não se equivalem porque ao tempo que o Estado busca o bem-estar social, as empresas buscam os lucros e pretendem a continuidade de seu serviço com a satisfação dos usuários, e a satisfação do usuário não equivale ao bem-estar da coletividade.

Em virtude das milícias digitais serem atividades coordenadas que se valem da demora da moderação para avançar e potencializar efeitos e, como os Estados não têm capacidade técnica para fazer controle de conteúdo online em tempo real, às empresas gestoras dos fluxos comunicacionais online foram cooptadas para realizar esse contingenciamento, ainda que em caráter emergencial, no intuito de minimizar os danos da propagação e restabelecer a ordem informacional ou impedir a articulação de ataques cibernéticos simultâneos ao redor do mundo.

Diante disso, operam dentro de seus limites particulares, sem a necessidade de ordem judicial prévia ou notificação individual para cada usuário que sofreu limitações, e como é medida de controle sistêmica, a revisão da filtragem feita pelas próprias empresas pode ser submetida ao âmbito judicial em virtude do interesse coletivo da sociedade ou ser analisada judicialmente de maneira individual para corrigir as distorções.

Na apreciação judicial, a ação seria provocada por quem se sentisse prejudicado e o prazo para a propositura da ação não seria a partir do conhecimento da exclusão de fake news disseminadas por milícias digitais, mas a partir da data do contingenciamento. Assim, não haveria a obrigatoriedade de comunicar a cada usuário em particular sobre as exclusões de fake

news disseminadas por milícias digitais operadas de forma sistêmica, mas haveria um prazo razoável para que o usuário que se sentisse prejudicado interpusesse ação judicial ou procedesse uma reclamação internamente. Apesar de entender que não haver a obrigação da empresa de comunicar individualmente cada usuário, subsistiria o dever de resposta à reclamação do usuário em prazo razoável.

Desse modo, as preocupações estão mais voltadas para a audiência do que para o indivíduo que propaga fake news em associação com atividade miliciana digital, caso a caso pode ser apurada a culpabilidade do perfil perpetrador ou restabelecido o direito de quem sofreu uma limitação indevida da liberdade de expressão, porque esse controle sistêmico, em tempo real e em escala que associa 02 critérios de moderação: comportamento e conteúdo, diminui a possibilidade de falhas nas filtragens e torna menos improvável a alegação, pelos perpetradores, de compartilhamento desavisado ou de erros de checagem sem os cuidados do homem médio, razão pela qual a audiência deve ser prestigiada e protegida em relação a questionável liberdade de expressão do usuário-perpetrador.

Em aspirações ideais, tecnologia e direito precisam interagir para que as medidas de controle de conteúdo online sejam democráticas, revestidas de legalidade, com as possibilidades operacionais do sistema claras, definidos e padronizados os parâmetros da atividade miliciana para que o enquadramento e controle não sejam arbitrários e as normas técnicas desenvolvidas pelo design para rastrear e limitar a disseminação de fake news por milícias sejam mais uniformes, no entanto, a realidade pragmática está muito distante dessas aspirações.

Desse modo, a regulação por lei mais abrangente associada a regulamentação por norma técnica que especificará os controles de acordo com as possibilidades tecnológicas tende a ser a estrutura jurídica desejável para moderação de conteúdo online segundo os procedimentos democráticos. As hipóteses de controle de conteúdo, caso a caso, dentro de cada empresa de gestão dos fluxos comunicacionais online promovem a fragmentação do sistema, com uma infinidade de comandos particulares que impactam o comportamento de milhões de pessoas.

Diante disso, as propostas legislativas estão em andamento e enquanto não aprovadas por instâncias democráticas, a filtragem algorítmica de fake news disseminadas por milícias digitais pode ser avaliada segundo o justo motivo que fundamentou a decisão de moderação ainda que de modo sistêmico, sopesando as limitações e danos com as medidas de contingenciamento abrangente para todos os usuários.

REFERÊNCIAS

- ABRUSIO, Juliana; MEDEIROS, Thamara. Fake News como instrumento de estratégia militar. *In: RAIS, Diogo (coord.) Fake News: a conexão entre a desinformação e o direito*. 2. ed. rev., atual. e ampl. São Paulo: Thomson Reuters Brasil, 2020.
- ALBRIGHT, Jonathan. Welcome to the era of fake news. *Media and Communication*, v. 5, n. 2, p. 87-89, 2017.
- ALMEIDA, Gilberto. Notas sobre utilização de inteligência artificial por agentes empresariais e suas implicações no âmbito do Direito do Consumidor. *In: FRAZÃO, Ana; MULHOLLAND, Caitlin (coord.) Inteligência Artificial e Direito: ética, regulação e responsabilidade*. São Paulo: Thompson Reuters Brasil, 2019.
- ARENT, Hanna. **Verdade e Política**. The New Yorker. Tradução: Manuel Alberto, fev. 1967.
- ARISTOTELES. **Política**. São Paulo: Martin Claret, 2017.
- ASH, Timothy. **Liberdade de Expressão: dez princípios para um mundo interligado**. Lisboa: Temas e Debates – Círculo de Leitores, 2017.
- BALDI, Vania. A construção viral da realidade: ciberpopulismos e polarização dos públicos em rede. *Observatorio (OBS*)*, n. 1, p. 4-20, 2018. Disponível no site <https://obs.obercom.pt/index.php/obs/article/view/1420/pdf> Acesso em: 14 set. 2021.
- BALKIN, Jack M. O Futuro da liberdade Expressão na Era Digital. *In: SAMPAIO, José Adércio Leite (Coord.) Liberdade de expressão no Século XXI*. Belo Horizonte: Del Rey, 2016.
- BALKIN, Jack M. Free speech is a triangle. *Columbia Law Review*, 2018. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3186205. Acesso em: 18 nov. 2022.
- BALKIN, Jack. How to Regulate (and Not Regulate) Social Media. *Academic Commons*, 2020. Disponível em: <https://academiccommons.columbia.edu/doi/10.7916/d8-jpb9-mr58>. Acesso em: 18 nov. 2022.
- BARLOW, J. P. **Declaração de Independência do Ciberespaço**. Suíça, 8 fev. 1996, dhnet@dhnet.org.br. Disponível em: <http://www.dhnet.org.br/ciber/textos/barlow.htm>. Acesso em: 4 nov. 2021.
- BARROSO, Luis Roberto. **Revolução tecnológica, crise da democracia e Constituição: direitos e políticas públicas num mundo em transformação**. Belo Horizonte: Fórum, 2021.
- BERARDI, Franco. **Depois do Futuro**. São Paulo: Ubu Editora, 2019.
- BINENBOJM, Gustavo; PEREIRA, Caio Mario da Silva. *In: FISS, Owen. A ironia da liberdade de expressão: estado, regulação e diversidade na esfera pública*. Rio de Janeiro: Renovar, 2005.

BOBBIO, Norberto. **O futuro da democracia. Uma defesa das regras do jogo**. Tradução Marco Aurélio Nogueira. 16ª edição. Rio de Janeiro/São Paulo: Paz e Terra, 2019.

BOBBIO, Norberto. **Era dos direitos**. Tradução Carlos Nelson Coutinho; apresentação de Celso Lafer. [Reimpr]. Rio de Janeiro: LTC, 2020.

BRANCO, Paulo Gustavo Gonet; MENDES, Gilmar Ferreira. **Curso de Direito Constitucional-Série IDP-2020**. 15 ed., São Paulo: Saraiva Educação SA, 2020.

BRASIL. Ministério da Saúde. **Aberta consulta pública sobre vacinação de crianças contra Covid-19**. Brasília: Ministério da Saúde, 2021a. Disponível em: <https://www.gov.br/saude/pt-br/assuntos/noticias/2021-1/dezembro/aberta-consulta-publica-sobre-vacinacao-de-criancas-contr-covid-19>. Acesso em: 18 maio 2022.

BRASIL. Supremo Tribunal Federal. **ADP 130**, Relator Ministro Carlos Ayres Britto. Tribunal Pleno, j. 30-04-2009. Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=AC&docID=605411>. Acesso em: 13 nov. 2021.

BRASIL. **Congresso Nacional. Lei nº 14.197, de 1º de setembro de 2021**. Acrescenta o Título XII na Parte Especial do Decreto-Lei nº 2.848, de 7 de dezembro de 1940 (Código Penal), relativo aos crimes contra o Estado Democrático de Direito; e revoga a Lei nº 7.170, de 14 de dezembro de 1983 (Lei de Segurança Nacional), e dispositivo do Decreto-Lei nº 3.688, de 3 de outubro de 1941 (Lei das Contravenções Penais). Brasília: Congresso Nacional, 2021b. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2019-2022/2021/Msg/VEP/VEP-427.htm Acesso em: 24 mar. 2023.

BRASIL. **Congresso Nacional. Lei nº 13.709, de 14 de agosto de 2018**, Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília: Congresso Nacional, 2018a. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/113709.htm. Acesso em: 20 mar. 2023.

BRASIL. **Resolução nº 332 do CNJ**, de 2020. Brasília: CNJ, 2020. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3429>. Acesso em: 19 maio 2022.

BRASIL. Supremo Tribunal Federal. Arguição de descumprimento de preceito fundamental 130 distrito federal. Relator: Min. Carlos Britto, Argte.(s) : Partido Democrático Trabalhista - PDT adv.(a/s) : Miro Teixeira e outro(a/s), Argdo.(a/s): Presidente Da República Adv.(a/s) : Advogado-Geral Da União Argdo.(a/s): Congresso Nacional Intdo.(a/s): Federação Nacional Dos Jornalistas Profissionais - FENAJ adv.(a/s): Claudismar Zupiroli E Outro(A/S), Intdo.(a/s): Associação Brasileira De Imprensa - Abi Adv.(a/s) : Thiago Bottino Do Amaral Intdo.(a/s): Artigo 19 Brasil Adv.(a/s) : Eduardo Pannunzio e outros. **DJE**, Brasília-DF, n. 208, 2009. Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=AC&docID=605411>. Acesso em: 20 mar. 2023.

BRASIL. **REsp 1642997 / RJ**, Recurso Especial 2016/0272263-4, Relatora Ministra Nancy Andrighi, T3 - Terceira Turma Julgado em 12/09/2017 Publicado no DJe em 15/09/2017. Disponível em: <https://processo.stj.jus.br/SCON/pesquisar.jsp>. Acesso em: 19 maio 2022.

BRASIL. Tribunal Superior Eleitoral. **Portaria nº 949, de 7 de dezembro de 2017**. Institui o Conselho Consultivo sobre Internet e Eleições. Brasília-DF: TSE, 2017a. Disponível em: <https://www.tse.jus.br/legislacao/compilada/prt/2017/portaria-no-949-de-7-de-dezembro-de-2017>. Acesso em: 20 mar. 2023.

BRASIL. Tribunal Superior Eleitoral. **Resolução nº 23.551, de 18 de dezembro de 2017, Dispõe sobre propaganda eleitoral, utilização e geração do horário gratuito e condutas ilícitas em campanha eleitoral nas eleições**. Brasília-DF: TSE, 2017b. Disponível em: <https://www.tse.jus.br/legislacao/compilada/res/2017/resolucao-no-23-551-de-18-de-dezembro-de-2017>. Acesso em: 20 mar. 2023.

BRASIL. Tribunal Superior Eleitoral. **Ação direta de inconstitucionalidade 4.451 Distrito Federal**. Brasília-DF: TSE, 2018b. Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=TP&docID=749287337>. Acesso em: 20 mar. 2023.

BRASIL. Câmara dos Deputados. **Projeto de Lei 2630/2020**, Institui a Lei Brasileira de Liberdade, Responsabilidade e Transparência na Internet. Brasília: Câmara dos Deputados, 2020a. Disponível em: <https://www.camara.leg.br/propostas-legislativas/2256735>. Acesso em: 20 mar. 2023.

BRASIL. Tribunal Superior Eleitoral. **Arguição de descumprimento de preceito fundamental 572**, relator: Min. Edson Fachin, repte.(s) :Rede Sustentabilidade, Adv.(a/s): Kamila Rodrigues Rosenda, adv.(a/s): Filipe Torri da Rosa; intdo.(a/s): Presidente do Supremo Tribunal Federal. Brasília-DF: TSE, 2020b. Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=TP&docID=754371407>. Acesso em: 20 mar. 2023.

BRASIL. Tribunal Superior Eleitoral. **Inquérito 4.781 Distrito Federal**, Relator: Min. Alexandre de Moraes, autor(a/s)(es) :sob sigilo, adv.(a/s) :sob sigilo. Brasília-DF: TSE, 2020c. Disponível em: <https://www.conjur.com.br/dl/decisao-alexandre-moraes-twitter.pdf>. Acesso em: 20 de mar. 2023.

BRASIL. Tribunal Superior Eleitoral. **Fato ou boato**: Esclarecimentos sobre informações falsas. Brasília-DF: TSE, 2020d. Disponível em: <https://www.justicaeleitoral.jus.br/fato-ou-boato/#>. Acesso em: 05 mar. 2023.

BRASIL. Tribunal Superior Eleitoral. **Arguição De Descumprimento De Preceito Fundamental 403 Sergipe**. Brasília-DF: TSE, 2021c. Disponível em: <https://www.stf.jus.br/arquivo/cms/noticiaNoticiaStf/anexo/ADPF403voto.pdf>. Acesso em: 20 mar. 2023.

BRASIL. Tribunal Superior Eleitoral. **Conselho Consultivo sobre Internet e Eleições se reúne no TSE (atualizada)**. Brasília-DF: TSE, 2022a. Disponível em: <https://www.tse.jus.br/comunicacao/noticias/2017/Dezembro/conselho-consultivo-sobre-internet-e-eleicoes-se-reune-no-tse>. Acesso em: 03 mar. 2023.

BRASIL. Tribunal Superior Eleitoral. **TSE firma novas parcerias com entidades e empresas para combater notícias falsas**. Brasília-DF: TSE, 2022b. Disponível em:

<https://www.tse.jus.br/comunicacao/noticias/2018/Junho/tse-firma-novas-parcerias-com-entidades-e-empresas-para-combater-noticias-falsas>. Acesso em: 08 mar. 2023.

BRASIL. Tribunal Superior Eleitoral. **Fato ou Boato**: site da Justiça Eleitoral verifica informações e alerta contra notícias falsas. Brasília-DF: TSE, 2022c. Disponível em: <https://www.tse.jus.br/comunicacao/noticias/2022/Outubro/fato-ou-boato-site-da-justica-eleitoral-verifica-informacoes-e-alerta-contra-noticias-falsas>. Acesso em: 05 mar. 2023.

BRASIL. Tribunal Superior Eleitoral. **Ações do TSE de combate à desinformação são elogiadas pelo Observatório da Universidade Complutense de Madrid**. Brasília-DF: TSE, 2022d. Disponível em: <https://www.tse.jus.br/comunicacao/noticias/2022/Dezembro/acoes-do-tse-de-combate-a-desinformacao-sao-elogiadas-pelo-observatorio-da-universidade-complutense-de-madrid>. Acesso em: 05 mar. 2023.

BRASIL. Tribunal Superior Eleitoral. Resolução nº 23.714, Processo Administrativo Nº 0601570-94.2022.6.00.0000 - Brasília - Distrito Federal, Relator: Ministro Alexandre de Moraes Interessado: Tribunal Superior Eleitoral Dispõe sobre o enfrentamento à desinformação que atinja a integridade do processo eleitoral. **DJE**, Brasília-DF, Ano 2022 - n. 213, 24 out. 2022e. Disponível em: <https://sintse.tse.jus.br/documentos/2022/Out/24/diario-da-justica-eletronico-tse-edicao-eleitoral/resolucao-no-23-714-de-20-de-outubro-de-2022-dispoe-sobre-o-enfrentamento-a-desinformacao-que-atinja>. Acesso em: 02 mar. 2023.

BRASIL. Supremo Tribunal Federal. Tema 987 - Discussão sobre a constitucionalidade do art. 19 da Lei n. 12.965/2014 (Marco Civil da Internet) que determina a necessidade de prévia e específica ordem judicial de exclusão de conteúdo para a responsabilização civil de provedor de internet, websites e gestores de aplicativos de redes sociais por danos decorrentes de atos ilícitos praticados por terceiros. Relator(a): Min. Dias Toffoli, Leading Case: RE 1037396. **DJE**, Brasília-DF, 2023a. Disponível em: <https://portal.stf.jus.br/jurisprudenciaRepercussao/verAndamentoProcesso.asp?incidente=5160549&numeroProcesso=1037396&classeProcesso=RE&numeroTema=987>. Acesso em: 16 mar. 2023.

BRASIL. Tribunal Superior Eleitoral. **JE em Ação**: enfrentamento da desinformação é trabalho constante desde 2017. Brasília-DF: TSE, 2023b. Disponível em: <https://www.tse.jus.br/comunicacao/noticias/2023/Marco/je-em-acao-enfrentamento-a-desinformacao-e-trabalho-constante-desde-2017>. Acesso em: 08 mar. 2023.

BRASIL. **Congresso Nacional. Lei nº 14.533, de 11 de janeiro de 2023**. Institui a Política Nacional de Educação Digital e altera as Leis nºs 9.394, de 20 de dezembro de 1996 (Lei de Diretrizes e Bases da Educação Nacional), 9.448, de 14 de março de 1997, 10.260, de 12 de julho de 2001, e 10.753, de 30 de outubro de 2003. Brasília: Congresso Nacional, 2023c. Disponível em: https://www.planalto.gov.br/ccivil_03/_Ato2023-2026/2023/Lei/L14533.htm#:~:text=LEI%20N%C2%BA%2014.533%2C%20DE%2011%20DE%20JANEIRO%20DE%202023&text=Institui%20a%20Pol%C3%ADtica%20Nacional%20de,30%20de%20outubro%20de%202003. Acesso em: 24 mar. 2023.

BUCCI, Eugênio. **Existe democracia sem verdade factual?** Organizado por Lucia Santella. Barueri, SP: Estação das Letras e Cores, 2019.

CALLEJÓN, Francisco Balaguer. Las dos grandes crisis del constitucionalismo frente a la globalización en el siglo XXI. **Revista de derecho constitucional europeo**, n. 30, p. 2, 2018.
 CALLEJÓN, Francisco Balaguer. Redes sociais, companhias tecnológicas e democracia. **REI – Revista Estudos Institucionais**, v. 6, n. 2, p. 579-599, 2020.

CALLEJÓN, Francisco Balaguer. O impacto dos novos mediadores da era digital na liberdade de expressão. **Espaço Jurídico Journal of Law [EJLL]**, v. 23, n. 1, p. 179-204, 2022a.

CALLEJÓN, Francisco Balaguer. **La constitución del algoritmo**. Zaragoza, Fundación Manuel Giménez Abad, 2022b.

CAMPOS, Ricardo; MARANHÃO, Juliano. Fake News e a autorregulação regulada nas redes sociais no Brasil: fundamentos constitucionais. *In*: ABOUD, Georges; NERY, Nelson Jr.; CAMPOS, Ricardo (org.). **Fake news e regulação**. São Paulo: Thomson Reuters; Revista do Tribunais, 2020.

CASTELLS, Manuel. **A Sociedade em Rede – A Era da Informação: Economia, Sociedade e Cultura**. v.1. São Paulo: Paz e Terra, 2020.

CERON, Wilson. Bots “do bem”, bots “do mal” e a importância da regulação. **Desinformante**, Coluna Ponto de Vista, 9 maio 2022. Disponível em: <https://desinformante.com.br/bots-do-bem-bots-do-mal-e-a-importancia-da-regulacao/>. Acesso em: 17 jul. 2022.

CHUDAKOV, Barry *in* RAINIE, Lee; ANDERSON, Janna; ALBRIGHT, Jonathan. **The Future of Free Speech, Trolls, Anonymity and Fake News Online**, 2017. Disponível em: <https://apo.org.au/node/75781>. Acesso em: 19 jul. 2022.

CUEVA, Ricardo Villas Boas. Alternativas para a remoção de fake news das redes sociais. *In*: ABOUD, Georges; NERY, Nelson Jr.; CAMPOS, Ricardo (org.). **Fake news e regulação**. São Paulo: Thomson Reuters; Revista do Tribunais, 2020.

DATASENADO, Instituto. Pesquisa: **Mais de 80% dos brasileiros acreditam que redes sociais influenciam muito a opinião das pessoas**. Disponível em: www12.senado.leg.br. Acesso em: 21 abr. 2021.

DOUEK, Evelyn. Content Moderation as Administration. **HARV. L. REV.**, v. 136, p. 1, 15-16, 2022. Disponível em <https://ssrn.com/abstract=4005326> or <http://dx.doi.org/10.2139/ssrn.4005326>. Acesso em: 02 ago. 2022.

ELIS-PETERSEN, Hannah. Facebook admits failings over incitement to violence in Myanmar. **The Guardian**, 6 nov. 2018. Disponível em: <https://www.theguardian.com/technology/2018/nov/06/facebook-admits-it-has-not-done-enough-to-quell-hate-in-myanmar>. Acesso em: 18 maio 2022,

FACEBOOK. Combating Misinformation. **Meta**, 2021. Disponível em: <https://about.fb.com/news/tag/misinformation/>. Acesso em: 02 de jan. 2022

FACEBOOK. Promover segurança e expressão. **Meta**, [202-]. Disponível em: <https://about.facebook.com/br/actions/promoting-safety-and-expression/>. Acesso em: 30 abr. 2022.

FACEBOOK. Transparency Center. **Meta**, 2022b. Disponível em: <https://transparency.fb.com/pt-br>. Acesso em: 15 fev. 2022.

FARIAS, Edilsom Pereira. **Colisão de direitos**: a honra, a intimidade, a vida privada e a imagem versus liberdade de expressão e comunicação. Porto Alegre: Sergio Antonio Fabris Ed., 2008.

FERNANDES, Victor Oliveira; MENDES, Gilmar Ferreira; FERNANDES, Victor Oliveira. Constitucionalismo digital e jurisdição constitucional: uma agenda de pesquisa para o caso brasileiro. **Revista Brasileira de Direito**, Passo Fundo, v. 16, n. 1, p. 1-33, jan./abr. 2020.

FERNANDES, Tarsila Ribeiro Marques. Democracia defensiva: origens, conceito e aplicação prática. **Revista de informação legislativa**, Brasília, V. 58, n. 230, p. 133-147, abr/jun. 2021. Disponível em: https://www12.senado.leg.br/ril/edicoes/58/230/ril_v58_n230_p133.pdf. Acesso em: 21 maio 2022.

FESTUGATTO, Adriana Martins Ferreira; MORAIS, José Luis Bolzan. **A democracia desinformada**: eleições e fake news. Porto Alegre: Livraria do Advogado, 2021.

FLEW, Terry; MARTIN, Fiona; SUZOR, Nicolau. Internet regulation as media policy: Rethinking the question of digital communication platform governance. **Journal of Digital Media & Policy**, v. 10, n. 1, p. 33-50, 2019.

FRANKS, Mary Anne. Unwilling avatars: Idealism and discrimination in cyberspace. **Colum. J. Gender & L.**, v. 20, p. 224-261, 2011. Disponível em https://repository.law.miami.edu/cgi/viewcontent.cgi?article=1304&context=fac_articles. Acesso em: 19 jul. 2022.

FREDES, Andrei Ferreira; MOLINARO, Carlos Alberto. Liberdade de Expressão e configuração do Ambiente Virtual: o controle do fluxo de informação e expressão na internet. In: HÄBERLE, Peter; MENDES, Gilmar Ferreira; BALAGUER CALLEJÓN, Francisco; SARLET, Ingo Wolfgang; STRAPAZZON, Carlos Luiz; AGUILAR CALAHORRO, Augusto. **Direitos fundamentais, desenvolvimento e crise do constitucionalismo multinível**. Porto Alegre: Fundação Fênix, 2020. (Livro em Homenagem a Jörg Luther)

FREITAS, Juarez; FREITAS, Thomas Bellini. **Direito e Inteligência artificial**: em defesa do humano. Belo Horizonte: Fórum, 2021.

FRYDMAN, Benoit. **O fim do Estado de Direito**: governar por standards e indicadores. Tradução Mara Beatriz Krug; revisão Jânia Maria Loes Saldanha. 2. ed. rev. Porto Alegre: Livraria do Advogado, 2018.

GIACCHETTA, André Zonaro. Atuação e Responsabilidade dos provedores diante das fake News e da desinformação. In: RAIS, Diogo (Coord.). **Fake News**: a conexão entre a

desinformação e o direito. 2 ed. rev., atual. e ampl. São Paulo: Thomson Reuters Brasil, 2020.

GRAGNANI, Juliana. Como identificar os diferentes tipos de fakes e robôs que atuam nas redes. **BBC**, 2017. Disponível em: <https://www.bbc.com/portuguese/brasil-42172154>. Acesso em: 22 maio 2022.

GROSS, Clarisse Piterman. Fake News e Democracia: discutindo o status do normativo falso e a liberdade de expressão. *In*: RAIS, Diogo (Coord.). **Fake News: a conexão entre a desinformação e o direito**. 2 ed. rev., atual. e ampl. São Paulo: Thomson Reuters Brasil, 2020.

GUNDIM, Wagner Wilson Deiró e PEREIRA, Flávio Leão Bastos. Fake News como Instrumento de Estratégia Militar. *In*: RAIS, Diogo (Coord.). **Fake News: a conexão entre a desinformação e o direito**. 2 ed. rev., atual. e ampl. São Paulo: Thomson Reuters Brasil, 2020.

GUTIERREZ, Andriei. É possível confiar em um sistema de Inteligência Artificial? Práticas em torno da melhoria da sua confiança, segurança e evidências de accountability. *In*: FRAZÃO, Ana; MULHOLLAND, Caitlin (coord.) **Inteligência Artificial e Direito: ética, regulação e responsabilidade**. São Paulo: Thompson Reuters Brasil, 2019.

HABERMAS, Jürgen. **A Inclusão do Outro: estudos de teoria política**. São Paulo: UNESP, 2018.

HABERMAS, Jürgen. **Teoria do Agir Comunicativo 1: racionalidade da ação e racionalidade social**. São Paulo: WMF Martins Fontes, 2019.

HABERMAS, Jürgen. **Facticidade e validade: contribuições para uma teoria discursiva do direito e da democracia**. Unesp, 2020.

HAN, Byung-Chul. **No enxame: perspectivas do digital**. Tradução de Lucas Machado. 2ª Reimpressão. Petrópolis: Vozes, 2019.

HARARI, Yuval Noah. **Sapiens: uma breve história da humanidade**. Tradução Janaína Marcoantonio. 51. ed. Porto Alegre, RS: L&PM, 2020.

HARTMANN, Ivar Alberto Martins; SARLET, Ingo Wolfgang. Direitos Fundamentais e Direito Privado: a Proteção da Liberdade de Expressão nas Mídias Sociais. **RDU**, Porto Alegre, v. 16, n. 90, p. 85-108, nov./dez. 2019.

HOFFMANN-RIEM, Wolfgang. **Teoria geral do direito digital: transformação digital: desafios para o Direito**. Rio de Janeiro: Forense, 2021.

HOLMES, Stephen; SUNSTEIN, Cass R. **O custo dos direitos: por que a liberdade depende dos impostos**. São Paulo: WMF-Martins fontes, 2019.

HORWITZ, Jeff ; SEETHARAMAN, Deepa. **Facebook Executives Shut Down Efforts to Make the Site Less Divisive: The social-media giant internally studied how it polarizes users, then largely shelved the research**. **The Wall Street Journal**, 26 maio 2020. Disponível em:

https://www.wsj.com/articles/facebook-knows-it-encourages-division-top-executives-nixed-solutions-11590507499?reflink=share_mobilewebshare. Acesso em: 18 maio 2022.

INNERARITY, Daniel. **Política para perplexos: o fim das certezas**. Petrópolis, RJ: Vozes, 2021.

INSTAGRAM. Our Commitment to Lead the Fight Against Online Bullying. **Instagram**, 08, Jul. 2019. Disponível em: <https://about.instagram.com/blog/announcements/instagrams-commitment-to-lead-fight-against-online-bullying/?refsrc=about.facebook.com/actions/promoting-safety-and-expression/>. Acesso em: 20 fev. 2022.

INSTAGRAM. Proteções e Atualizações da Política sobre COVID-19 e Vacinas. **Instagram**, [2020]. Disponível em: https://help.instagram.com/697825587576762?helpref=faq_content. Acesso em: 20 fev. 2022.

INSTAGRAM. Termos de uso. **Instagram**, 2022b. Disponível em: <https://help.instagram.com/581066165581870>. Acesso em: 20 fev. 2022.

INSTAGRAM. Diferenças entre a NetzDG e as Diretrizes da Comunidade do Instagram. **Instagram**, 2022c. Disponível em: https://help.instagram.com/1787585044668150/?helpref=uf_share. Acesso em: 13 maio 2022.

INSTAGRAM. Termos e impressão. **Instagram**, Central de Ajuda, 4 jan. 2022a. Disponível em: https://help.instagram.com/581066165581870/?helpref=hc_fnav. Acesso em: 20 fev. 2022.

JESUS, Diego Santos Vieira; STEIBEL, Fabro; VICENTE, Victor Freitas. Possibilidades e Potenciais da Utilização da Inteligência Artificial. In: FRAZÃO, Ana; MULHOLLAND, Caitlin (coord.) **Inteligência Artificial e Direito: ética, regulação e responsabilidade**. São Paulo: Thompson Reuters Brasil, 2019.

KAKUTANI, Michiko. **A morte da verdade**. Tradução André Czarnobai e Marcela Duarte. 1. ed. Rio de Janeiro: Intrínseca, 2018.

KLONICK, Kate. The new governors: The people, rules, and processes governing online speech. **Harv. L. Rev.**, v. 131, p. 1598, 2017. Disponível em https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2937985. Acesso em: 21 jan. 2021.

KLONICK, Kate. Por dentro da equipe do Facebook que lidou com o tiroteio em Christchurch. **The New Yorker**, 25 abr. 2019. Disponível em: <https://www.newyorker.com/news/news-desk/inside-the-team-at-facebook-that-dealt-with-the-christchurch-shooting>. Acesso em: 08 jan. 2023.

LESSIG, Lawrence. Code is law: on liberty in cyberspace. **Harvard Magazine**, jan. 2000. Disponível em: <https://www.harvardmagazine.com/2000/01/code-is-law-html>. Acesso em: 12 nov. 2021.

LEVITSKY, Steven; ZIBLATT, Daniel. **Como as democracias morrem**. Rio de Janeiro: Zahar, 2018.

LEVY, Pierre. **Cibercultura**. 3. Ed., 3ª reimp. Tradutor Carlos Irineu da Costa. São Paulo: Editora 34, 2018.

LOBO, Edilene de Moraes; BOLZAN, José Luis; NEMER, David. DEMOCRACIA ALGORITMICA: O FUTURO DA DEMOCRACIA E O COMBATE ÀS MILÍCIAS DIGITAIS NO BRASIL. **Revista Culturas Jurídicas**, v. 7, n. 17, 2020.

LOSEKANN, Cristiana. A esfera pública habermasiana, seus principais críticos e as possibilidades do uso deste conceito no contexto brasileiro. **Pensamento Plural**, n. 4, p. 37-57, 2014.

LOVELUCK, Benjamin. **Redes, Liberdades e Controle**: uma genealogia política da internet. Tradução: Guilherme João de Freitas Teixeira. Petrópolis, RJ: Vozes, 2018.

MACEDO JR, Ronald Porto. Liberdade de Expressão: que lições podemos aprender com a experiência americana. *In*: FARIA, José Eduardo (Org.). **A Liberdade de Expressão e as Novas Mídias**. 1. ed. São Paulo: Perspectiva, 2020.

MAGRANI, Eduardo. **Democracia conectada**: a internet como ferramenta de engajamento político-democrático. Rio de Janeiro: FGV Direito Rio, 2014.

MAGRANI, Eduardo; OLIVEIRA, Renan Medeiros. A esfera pública (forjada) na era das fake news e dos filtros-bolha. **Cadernos Adenauer**, n. 4, 2018. Disponível em <http://eduardomagrani.com/wp-content/uploads/2019/05/PUBLICACAO-nova-2019-KA-Cadernos-2018.4-site.pdf>. Acesso em: 30 nov. 2022.

MANSUR, Rafael e SOARES, Felipe Ramos Ribas. A tese da posição preferencial da liberdade de expressão frente aos direitos da personalidade: análise crítica a luz da legalidade constitucional *In*: MORAES, Bruno Terras; SCHREIBER, Anderson; TEFFÉ, Chiara Spadaccini (coord.). **Direito e mídia**: tecnologia e liberdade de expressão. Indaiatuba, SP: Foco, 2020.

MART, Susi; ALEXY, Robert. **Proporcionalidade e Internet**. Porto Alegre: Sergio Antônio Fabris Ed, 2020.

MARTINS, Humberto. **Lei da transparência e sua aplicação na administração pública: valores, direito e tecnologia em evolução**. 2018. Disponível em: http://www.stj.jus.br/internet_ministros/Discursos/0001182/LEI%20DA%20TRANSPAR%20ANCIA%20E%20SUA%20APLICAC%3%87%20P%20ABICA%20VALORES,%20DIREITO%20E%20TECNOLOGIA%20EM%20EVOLUC%3%83O.pdf. Acesso em: 03 nov 2022.

MEDRADO, Vitor Amaral. **A Liberdade de Expressão e a Justiça Brasileira**: tolerância, discurso e de ódio e democracia. 2. ed., rev. e ampl. Belo Horizonte: Dialética, 2019.

MENDES, Laura Schertel Ferreira. Autodeterminação informativa: a história de um conceito. **Pensar-Revista de Ciências Jurídicas**, v. 25, n. 4, 2020. Disponível em: <https://scholar.google.com.br/scholar?hl=pt->

BR&as_sdt=0%2C5&q=laura+schertel+mendes+autodetermina%C3%A7%C3%A3o&oq=laura+schertel+mendes+autodetermin. Acesso em: 14 out. 2022.

MENEZES, Paulo Brasil. **Fake news: modernidade, metodologia e regulação**. Salvador: Jus Podium, 2020.

META. Fake Accounts. **Meta, 2022a**. Disponível em: <https://transparency.fb.com/data/community-standards-enforcement/fake-accounts/facebook/>. Acesso em: 10 maio 2022.

META. Facebook e Instagram terão canal dedicado ao TSE para denúncias. **Meta, 2022**. Disponível em: <https://about.fb.com/br/news/2022/02/facebook-e-instagram-terao-canal-dedicado-ao-tse-para-denuncias/>. Acesso em: 10 maio 2022.

META. Community Standards Enforcement Report. **Meta, 2023a**. Disponível em: <https://transparency.fb.com/data/community-standards-enforcement/?source=https%3A%2F%2Ftransparency.facebook.com%2Fcommunity-standards-enforcement>. Acesso em: 10 maio 2022.

META. Coordinated Inauthentic Behavior Explained. **Meta, 2018**. Disponível em <https://about.fb.com/news/2018/12/inside-feed-coordinated-inauthentic-behavior/>. Acesso em: 10 maio 2022.

MILL, John Stuart. **Sobre a liberdade**. Tradução de Maria Aparecida Sargiolatto. Campinas, SP: Vide Editorial, 2018.

MORELLI, Alessandro; POLLICINO, Oreste. Metaphors, Judicial Frames, and Fundamental Rights in Cyberspace. **The American Journal of Comparative Law**, v. 68, n. 3, p. 616-646, 2020.

MOROZOV, Evgeny. **Big Tech: a ascensão dos dados à morte da política**. 2ª reimpressão. São Paulo: Ubu Editora, 2020.

NITRINI, Rodrigo Vidal. **Liberdade Expressão nas redes sociais: o problema jurídico da remoção de conteúdo pelas plataformas**. Belo Horizonte: Dialética, 2021.

NÓBREGA, Liz. Brasileiros confiam mais nas notícias via Google e WhatsApp que na mídia tradicional. **Desinformante: informação confiável sobre desinformação**, [s. l.], 27, set., 2022. Disponível em: <https://desinformante.com.br/brasileiros-confiam-mais-nas-noticias-via-google-e-whatsapp-que-na-midia-tradicional/>. Acesso em: 15 mar. 2023.

NOELLE-NEUMANN, Elisabeth. La espiral del silencio. La opinión pública y los efectos de los medios de comunicación. **Communication & Society**, v. 6, n. 1-2, p. 9-28, 1993. Disponível em <https://dadun.unav.edu/bitstream/10171/8031/1/20091113225431.pdf>. Acesso em: 22 set. 2022.

NOHARA, Irene Patrícia. Desafios da Ciberdemocracia Diante do Fenômeno das Fake News: regulação estatal em face dos perigos da desinformação. *In*: RAIS, Diogo (coord.). **Fake News: a conexão entre a desinformação e o direito**. 2. ed. rev., atual. e ampl. São Paulo: Thomson Reuters Brasil, 2020.

NORONHA, José Otávio de. Os avanços da inteligência artificial no STJ: caminho para uma Justiça mais rápida e efetiva. *In*: SILVA, Rodrigo da Cunha Silva; TEPEDINO, Gustavo (coord.). **O Direito Civil da era da inteligência artificial**. São Paulo: Revista dos Tribunais, 2020.

OLIVEIRA, Carlos Affonso Pereira; SOUZA, Jordan Vinícius. Sobre os ombros de robôs? A inteligência artificial sobre fascínios e desilusões. *In*: FRAZÃO, Ana; MULHOLLAND, Caitlin (coord.) **Inteligência Artificial e Direito: ética, regulação e responsabilidade**. São Paulo: Thompson Reuters Brasil, 2019.

O'NEIL, Cathy. **Algoritmos de destruição em massa**. Tradução Rafael Abraham. 1. ed. Santo André: Rua do Sabão, 2020.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS. **Report of the Working Group on Internet Governance. Bogis-Bossey**; 2005. p. 1 - 24. (05.41622). p. 4. Disponível em: <https://www.wgig.org/docs/WGIGREPORT.pdf>. Acesso em: 10 maio 2022.

PINCH, Trevor J.; BIJKER, Wiebe E. The social construction of facts and artefacts: or how the Sociology of Science and the Sociology of Technology might benefit each other. **Social Studies of Science**, Sage, London, Beverly Hills and New Delhi, v. 14, p. 399-441, 1984. Disponível em: <http://sss.sagepub.com/cgi/content/abstract/14/3/399>. Acesso em: 8 nov. 2021

PINHEIRO, Joel. Fake News e o futuro da nossa civilização. *In*: BARBOSA, Mariana (Org.). **Pós Verdade e Fake News: reflexões sobre a guerra de narrativas**. 1. ed. Rio de Janeiro: Cobogó, 2019.

PITRUZZELLA, Giuseppe. La libertà di informazione nell'era di Internet. **Media Laws (1)**, p. 19-47, 2018. Disponível em <https://www.medialaws.eu/wp-content/uploads/2019/05/1.-Pitruzzella.pdf>. Acesso em: 19 jul. 2022.

QUILES, Marcos G. **A dualidade da inteligência artificial no combate às fake News**. 2022. Disponível em: <http://desiformante.com.br/a-dualidade-da-inteligencia-artificial-no-combate-as-fake-news/>. Acesso em: 8 nov 2022.

RAIS, Diego. Desinformação no contexto democrático. *In*: ABBOUD, Georges; NERY, Nelson Jr.; CAMPOS, Ricardo (org.). **Fake news e regulação**. São Paulo: Thomson Reuters; Revista do Tribunais, 2020.

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

ROCILLO, Paloma; DUARTE, Felipe Marcelo; ROMAN, Juliana. O que é governança da internet? uma revisão sistemática de literatura. **Revista Jurídica da UFRSA**, v. 5, n. 10, p. 151-178, 2021.

SARLET, Ingo Wolfgang; MOLINARO, Carlos Alberto. Liberdade de expressão!(Superando os limites do'politicamente (in) correto'. **Revista da AJURIS**, 2012.

SCHREIBER, Anderson. Liberdade de Expressão e Tecnologia. In: MORAES, Bruno Terras; SCHREIBER, Anderson; TEFFÉ, Chiara Spadaccini (coord.). **Direito e mídia: tecnologia e liberdade de expressão**. Indaiatuba, SP: Foco, 2020.

SCHAUER, Frederick. A. Liberdade Expressão em Decisões na Europa e nos Estados Unidos: um estudo de caso na arquitetura constitucional comparada. In: SAMPAIO, José Adércio Leite (coord.). **Liberdade de expressão no Século XXI**. Belo Horizonte: Del Rey, 2016.

SEDLER, Robert. A. Um ensaio sobre a Liberdade Expressão: Estados Unidos versus o resto do mundo. In: SAMPAIO, José Adércio Leite (coord.). **Liberdade de expressão no Século XXI**. Belo Horizonte: Del Rey, 2016.

SEMINÁRIO INTERNACIONAL FAKE NEWS E ELEIÇÕES (ABERTURA). Publicado pelo canal da Justiça Eleitoral. 2020. 1 vídeo (1h27min19s). Disponível em: <https://www.youtube.com/live/8-O5fL5Pt9w?feature=share>. Acesso em: 04 mar. 2023.

SILVEIRA, Sérgio Amadeu. A noção de modulação e os sistemas de algoritmos. In: AVELINO, Rodolfo; SILVEIRA, Sérgio Amadeu; SOUZA, Joyce (org). **A Sociedade do Controle: Manipulação e Modulação das Redes Digitais**. São Paulo: Hedra, 2018.

SOUZA, Carlos Afonso e TEFFE, Chiara Spaccini. Fake News e eleições: identificando e combatendo a desordem informacional. In: MORAES, Bruno Terras; SCHREIBER, Anderson; TEFFÉ, Chiara Spadaccini (coord.). **Direito e mídia: tecnologia e liberdade de expressão**. Indaiatuba, SP: Foco, 2020

SUZOR, Nicolas. Digital constitutionalism: Using the rule of law to evaluate the legitimacy of governance by platforms. **Social Media+ Society**, v. 4, n. 3, p. 2056305118787812, 2018. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2909889. Acesso em: 08 ago. 2022.

SUNSTEIN, Cass R. **A verdade sobre os boatos: como se espalham e por que acreditamos neles**. Rio de Janeiro: Elsevier, 2010.

SUNSTEIN, Cass R. **The cost-benefit revolution**. Cambridge, MA: MIT Press, 2018.

TEUBNER, Gunther. Horizontal effects of constitutional rights in the internet: A legal case on the digital constitution. **Italian LJ**, v. 3, p. 193, 2017. Disponível em: https://d1wqtxts1xzle7.cloudfront.net/54441161/teubner-libre.pdf?1505479837=&response-content-disposition=inline%3B+filename%3DG_Teubner_Horizontal_Effects_of_Constitu.pdf&Expires=1672851612&Signature=R1YaorTj61wagvkkCpO4IB8~T7sI8~-CbkuUDhslx57A6mtDTtNM3xsmLv4S-TaAfdJDo1cilo9iHWb~B6RDogv9NrkBJlhKu9Y6qzdkLzfr4MaZgy~Cb4yIDXPrzIIRXlhKVag1ZGplz484JrM72aR8X0QnVzsqbFYc7oI8zAsXIIdwKPPVadG86nY7cM4gWwlR24tegMS9AZMn2-uSNh2sH9Rz2p8uCFnabaTNi~tLpTaT9M1kMeBh31Iy9j5GTjH5c7ZiAd58IrsnGOZF1tDN

Y-7g5SVzS3C2HZjGdHYfaZQPUjiNt0KIwNwKNkkUqnNS4FeTFph1w--Qy7fg__&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA. Acesso em: 12 fev. 2022.

TWITTER. **Our range of enforcement options.** **Twitter**, Central de Ajuda, [20--]. Disponível em: <https://help.twitter.com/en/rules-and-policies/enforcement-options>. Acesso em: 05 maio 2022.

TWITTER. A nossa gama de opções de aplicação. **Twitter**, Central de Ajuda, [20--]. <https://help.twitter.com/en/rules-and-policies/enforcement-options>. Acesso em: 05 maio 2022

TWITTER. **Uma atualização sobre denúncias de potencial desinformação no Twitter.** 17 janeiro 2022b. Disponível em: https://blog.twitter.com/pt_br/topics/company/2022/uma-atualizacao-sobre-denuncias-de-potencial-desinformacao-no-tw. Acesso em 04 de maio de 2022.

UNIÃO EUROPEIA. **Carta Europeia de Ética Sobre O Uso da Inteligência Artificial em Sistemas Judiciais e Seu Ambiente.** Estrasburgo, 2018. Disponível em: https://rm.coe.int/carta-etica-traduzida-para-portugues-revista/168093b7e0#_Toc530141230. Acesso em: 19 maio 2022.

UNIÃO EUROPEIA. **Proposta de Regulamento do Parlamento Europeu e do Conselho.** Bruxelas, 2021. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:52021PC0206&from=EN>. Acesso em: 19 maio 2022.

UOL. **Chefe do Whatsapp diz que apoia PL das fake news.** Disponível em: <https://www.uol.com.br/tilt/noticias/redacao/2021/12/06/chefe-do-whatsapp-diz-que-apoia-pl-das-fake-news-discordamos-de-detalhes.amp.htm>. Acesso em 19 de maio de 2022.

VALENTE, Maria Giorgetti. A Liberdade de Expressão na Internet: da Utopia à Era das Plataformas. *In*: FARIA, José Eduardo (org.). **A liberdade de expressão e as novas mídias.** 1. ed. São Paulo: Editora Perspectiva, 2020.

VANNUCHI, Camilo. O direito à comunicação e os desafios da regulação dos meios no Brasil. **Galáxia (São Paulo)**, p. 167-180, 2018.

VEJA. **As milícias digitais que impulsionam o sucesso de estrelas como Anitta:** Grupos poderosos de fãs são capazes de dinamitar velhas certezas sobre o que determina ou não a fama. 2022. Disponível em: <https://veja.abril.com.br/cultura/as-milicias-digitais-que-impulsionam-o-sucesso-de-estrelas-como-anitta/>. Acesso em 30 de maio de 2022

VENTURINI, Tommaso. De notícias falsas a lixo eletrônico: a política de dados da viralidade online. *In*: **Política de Dados** . Routledge, 2019. pág. 123-144. Disponível no site https://scholar.google.com.br/scholar?hl=pt-BR&as_sdt=0%2C5&q=tommaso+venturini+From+Fake+to+Junk+News%2C+the+Data+Politics+of+Online+Virality&btnG= Acessado em 02 de setembro de 2021

WALDRON, Jeremy. **Law and disagreement.** New York, Oxford University Press Inc: 2004.

WHATSAPP. **Termos de Serviço do WhatsApp**. Disponível em: <https://www.whatsapp.com/legal/terms-of-service?eea=0#terms-of-service-availability-and-termination-of-our-services>. Acesso em 15 maio 2022.

WHATSAPP. Termos de Serviço do WhatsApp. **WhatsApp**, 4 jan. 2021. Disponível em: <https://www.whatsapp.com/legal/terms-of-service?eea=0#terms-of-service-acceptable-use-of-our-services>. Acesso em: Acesso em 15 maio 2022

WIELSCH, Dan. Os ordenamentos das redes: Termos e condições de uso – Códigos – Padrões da comunidade. *In*: ABOUD, Georges; NERY, Nelson Jr.; CAMPOS, Ricardo (org.). **Fake news e regulação**. São Paulo: Thomson Reuters/Revista do Tribunais, 2020.

WU, Tim. **Impérios da Comunicação: do telefone à internet, da A&T ao Google**. Tradução Cláudio Rio de Janeiro: Zahar, 2012.

ZUBOFF, Shoshana. **A Era do Capitalismo de Vigilância**: a luta por um futuro humano na fronteira do poder. Tradução George Schlesinger. Rio de Janeiro: Intrínseca, 2020.