



**Universidade do Estado do Rio de Janeiro**

Centro de Ciências Sociais

Faculdade de Direito

Bernardo José Oliveira Araujo

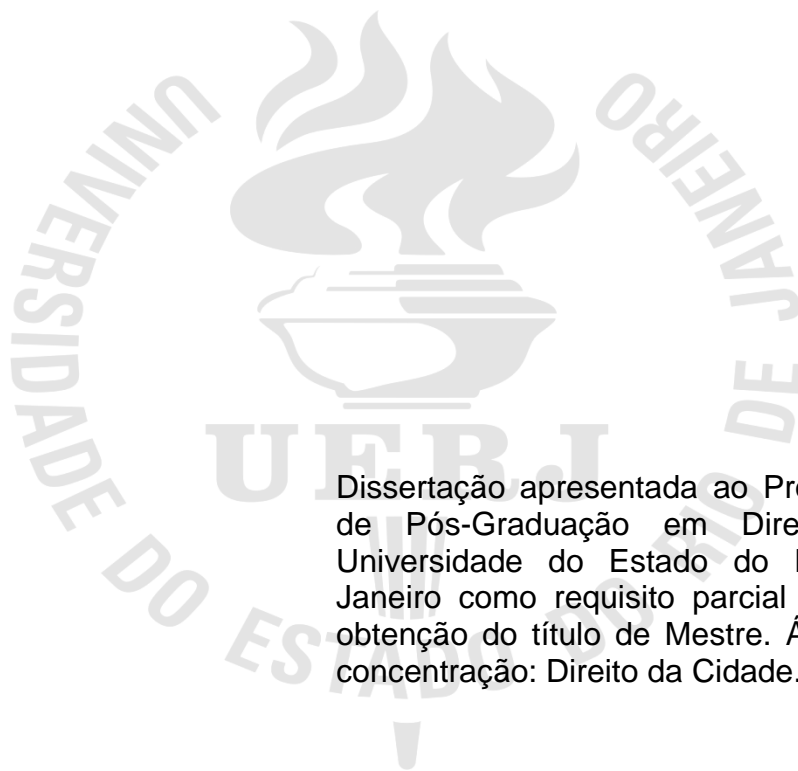
**Regulação das Redes Sociais:  
considerações sobre liberdade de expressão em Sociedades  
Algorítmicas**

Rio de Janeiro

2019

Bernardo José Oliveira Araujo

**Regulação das Redes Sociais:  
considerações sobre liberdade de expressão em Sociedades Algorítmicas**



Dissertação apresentada ao Programa de Pós-Graduação em Direito da Universidade do Estado do Rio de Janeiro como requisito parcial para a obtenção do título de Mestre. Área de concentração: Direito da Cidade.

Orientador: Prof. Dra. Vânia Siciliano Aieta

Rio de Janeiro

2019

CATALOGAÇÃO NA FONTE  
UERJ/REDE SIRIUS/BIBLIOTECA CCS/C

A663

Araújo, Bernardo José Oliveira.

Regulação das redes sociais: considerações sobre liberdade de expressão em sociedades algorítmicas / Bernardo José Oliveira Araújo. – 2019.

175f.

Orientador: Prof<sup>ª</sup>. Dra. Vânia Siciliano Aieta.

Dissertação (Mestrado). Universidade do Estado do Rio de Janeiro, Faculdade de Direito.

1. Regulação - Teses. 2.Redes sociais – Teses. 3. Liberdade de expressão – Teses. I.Aieta, Vânia Siciliano. II. Universidade do Estado do Rio de Janeiro. Faculdade de Direito. III. Título.

CDU 316.422

Bibliotecária: Marcela Rodrigues de Souza CRB7/5906

Autorizo, apenas para fins acadêmicos e científicos, a reprodução total ou parcial desta tese, desde que citada a fonte.

---

Assinatura

---

Data

Bernardo José Oliveira Araujo

**Regulação das Redes Sociais:  
considerações sobre liberdade de expressão em Sociedades Algorítmicas**

Dissertação apresentada ao Programa de Pós-Graduação em Direito da Universidade do Estado do Rio de Janeiro como requisito parcial para a obtenção do título de Mestre. Área de concentração: Direito da Cidade.

Aprovada em 26 de fevereiro de 2019.

Banca Examinadora:

---

Prof.<sup>a</sup> Dra. Vânia Siciliano Aieta (Orientadora)  
Faculdade de Direito – UERJ

---

Prof. Dr. Carlos Affonso Pereira de Souza  
Faculdade de Direito – UERJ

---

Prof. Dr. Marcelo Weick Pogliese  
Universidade Federal da Paraíba

Rio de Janeiro

2019

*Antes mundo era pequeno  
Porque Terra era grande  
Hoje mundo é muito grande  
Porque Terra é pequena  
Do tamanho da antena parabólicamará  
Ê, volta do mundo, camará  
Ê-ê, mundo dá volta, camará  
Antes longe era distante  
Perto, só quando dava  
Quando muito, ali defronte  
E o horizonte acabava  
Hoje lá trás dos montes, den de casa,  
camará  
Ê, volta do mundo, camará  
Ê-ê, mundo dá volta, camará*

*Trecho da música Parabolicamará de  
Gilberto Gil.*

## RESUMO

ARAUJO, Bernardo J. O. *Regulação das Redes Sociais: considerações sobre liberdade de expressão em Sociedades Algorítmicas*. 2019. 175 f. Dissertação (Mestrado em Direito) – Faculdade de Direito, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2019.

Mentiras são tão antigas quanto a história humana, mas a sua faceta moderna, mundialmente conhecida como *fake news*, representa um óbice maior para a Democracia nas Sociedades Algorítmicas. Nesse contexto, o trabalho almeja delinear as principais diferenças entre o ecossistema informacional na esfera pública clássica e na esfera pública digital, analisando o fenômeno da desinformação dentro de um grupo mais amplo de danos praticados nas redes sociais, tendo em vista que quaisquer regulações impostas entram igualmente em choque com a liberdade de expressão. A partir daí, sintetizando o cenário regulatório e a complexidade em regular setores disruptivos com incertezas científicas, busca-se inspiração na evolução jurisprudencial da *public forum doctrine* na Suprema Corte norte-americana para determinar quais tipos de restrições à fala são constitucionalmente permissíveis. O objetivo é contribuir com o debate regulatório, expondo algumas premissas interdisciplinares necessárias para pensarmos a regulação da liberdade de expressão nas redes sociais, especialmente no Brasil.

Palavras-chave: Regulação. Redes Sociais. Plataformas Interativas. Liberdade de Expressão. Democracia. Direito Administrativo

## ABSTRACT

ARAUJO, Bernardo J. O. *Social Media Regulation: considerations about free speech in Algorithmic Societies*. 2019. 175 f. Dissertação (Mestrado em Direito da Cidade) – Faculdade de Direito, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2019.

Lies are as old as human history, but its modern facet, world-wide known as fake news, represents a major obstacle to Democracy in Algorithmic Societies. In this context, the work aims to delineate the main differences between the informational ecosystem in the classical public sphere and in the digital public sphere, analyzing the phenomenon of disinformation within a broader group of online harms in social media, since any imposed regulations also shock with freedom of speech. Then, synthesizing the regulatory landscape and the complexity within regulating disruptive sectors with scientific uncertainties, we seek inspiration in the US Supreme Court's jurisprudential evolution of the public forum doctrine to determine what kinds of speech restrictions are constitutionally permissible. The goal is to contribute to the regulatory debate, presenting some necessities interdisciplinary premises to think about the regulation of freedom of expression in social media, especially in Brazil.

Keywords: Regulation. Social Media. Social Platforms. Freedom of Speech.  
Democracy. Administrative Law

## SUMÁRIO

	<b>INTRODUÇÃO.....</b>	<b>8</b>
<b>1</b>	<b>O ESVAZIAMENTO DO ECOSISTEMA INFORMACIONAL: COMO É PRODUZIDA A DESINFORMAÇÃO?.....</b>	<b>14</b>
1.1	<b>Pressupostos teóricos: sobre a arquitetura de controle das redes sociais.....</b>	<b>18</b>
1.1.1	<u>As redes sociais como canais de distribuição de conteúdo informacional: filtros-bolha e cascatas cibernéticas: polarização e extremismo.....</u>	<b>25</b>
1.1.2	<u>Micro-targeting e o caso Cambridge Analytica.....</u>	<b>34</b>
1.1.3	<u>O uso de robôs e os modelos de negócio online.....</u>	<b>37</b>
1.2	<b>O modelo russo de propaganda: guerra cibernética via desinformação: troll armies e overload informacional.....</b>	<b>43</b>
1.2.1	<u>O grande volume de informação: múltiplos veículos e canais de comunicação.....</u>	<b>48</b>
1.2.2	<u>O modelo é rápido, contínuo e repetitivo.....</u>	<b>52</b>
1.2.3	<u>O modelo não tem compromisso com a realidade objetiva.....</u>	<b>54</b>
1.2.4	<u>Falta de compromisso com a consistência jornalística.....</u>	<b>58</b>
<b>2</b>	<b>PENSANDO UM MODELO DE ESFERA PÚBLICA DIGITAL.....</b>	<b>61</b>
2.1	<b>A esfera pública clássica e a esfera pública digital: semelhanças e diferenças.....</b>	<b>65</b>
2.2	<b>A ascensão do amadorismo e as novas características da comunicação.....</b>	<b>78</b>



3	<b>UM PANORAMA REGULATÓRIO DAS INTERMEDIÁRIAS DA INTERNET: <i>OLD SCHOOL</i> VERSUS <i>NEW SCHOOL</i> SPEECH REGULATION</b> .....	82
3.1	Quais danos buscamos evitar e quais os principais desafios ao regular tecnologias como as plataformas interativas?.....	98
3.2	O princípio da precaução e o dever de cuidado como nortes regulatórios: é possível exigir um dever de monitoramento das plataformas?.....	105
4	<b>SOBRE A <i>PUBLIC FORUM DOCTRINE</i>: O QUE PENSA A SUPREMA CORTE NORTE AMERICANA SOBRE A LIBERDADE DE EXPRESSÃO EM ESPAÇOS PÚBLICOS: NEUTRALIDADE, RAZOABILIDADE, MANUTENÇÃO DE OUTROS CANAIS DE EXPRESSÃO</b> .....	120
4.1	Devemos enxergar as redes sociais como fóruns públicos?.....	136
5	<b>A POSIÇÃO ALEMÃ: A <i>NETWORK ENFORCEMENT ACT</i> É UMA REGULAÇÃO BRANDA?</b> .....	144
5.1	Principais objeções à <i>NetzDG</i> : <i>overblocking</i> e aplicação privada do direito.....	147
5.2	A configuração legislativa da <i>NetzDG</i> e suas principais características.....	154
	<b>SÍNTESE OBJETIVA E CONCLUSÕES</b> .....	162
	<b>COMENTÁRIOS METODOLÓGICOS E PROCEDIMENTAIS</b> .....	167
	<b>REFERÊNCIAS</b> .....	171

## INTRODUÇÃO

A explosão das redes sociais como fontes e canais de informação produziu algumas das mais impactantes mudanças sociais das últimas décadas. Não está enganado quem diz que, nas democracias modernas, a ideia de esfera pública clássica foi ampliada e transposta para o mundo digital, havendo, de fato, uma esfera pública digital com um novo tipo de ecossistema informacional - ambos merecedores de atenção e estudo. Assim, embora já se saiba que o consumo de informação pelas redes sociais vem causando severas distorções nas democracias, o aproveitamento dessa arquitetura por atores políticos agravou o problema e evidenciou que as redes sociais, plataformas interativas já consideradas fundamentais para a disseminação de informação, também podem ser manuseadas como instrumento de controle extremamente eficaz.

É neste contexto que a problemática da desinformação vem se mostrando um dos principais temas a serem abordados, em especial pelo Direito Constitucional, Administrativo, Eleitoral, ou melhor, pelas disciplinas que tangenciam estudos sobre Democracia. Seja pelos casos de notícias falsas (*fake news*) em eleições ao redor do mundo, ou pelos já incontáveis acontecimentos de menor expressão, ou até mesmo pelo nascimento de legislações regulatórias, a única conclusão possível é que o ecossistema informacional e a esfera pública atravessam uma crise de escala global, exigindo cuidados específicos nas democracias conectadas.

Há de ser feita uma nota extremamente importante para a compreensão da linha mestra traçada neste trabalho, porque, embora o ponto de partida seja a análise do fenômeno da desinformação como um *harm* específico a que se visa combater nas redes sociais, o tema geral consiste no debate sobre as problemáticas envolvidas na tentativa de se regular tecnologias disruptivas como as redes sociais. Portanto, é crucial deixar claro que eventuais regulações para *fake news* acabam esbarrando nos mesmos problemas de soluções regulatórias pensadas para outros *online harms*. Em suma, regular as redes sociais sempre envolverá, em alguma medida, a restrição da liberdade de expressão na esfera pública e as dificuldades arquitetônicas da Internet. Nesse sentido, iniciamos o trabalho falando sobre as *fake news* e os mecanismos que as potencializam, porque, além de ser uma temática premente no atual contexto político, serve perfeitamente para se introduzir alguns dos principais debates sobre a regulação das redes sociais como um todo. Sendo assim, a partir de certo momento, inserimos a desinformação dentro de um grupo mais abrangente de *online harms*

justamente para direcionar o trabalho à questão mais abrangente: a regulação da liberdade de expressão nas redes sociais.

Partindo desta questão geral, e a partir da seleção de temas chaves e conexos, este trabalho busca primeiro analisar como é produzida a desinformação nas redes sociais, mídias essas que também chamaremos de plataformas interativas, e, em momento oportuno, de novas intermediárias da comunicação. Sendo assim, o debate político-jurídico central deste trabalho consiste em investigar se as mudanças estruturais na esfera pública e no ecossistema informacional, causadas pelo desenvolvimento tecnológico das Sociedades Algorítmicas, requerem (ou não) uma atuação regulatória estatal, com o fim de preservar a democracia e a liberdade de expressão, fins estes que estão entrando em paradoxal choque na atualidade.

Diante desse tema central, desdobram-se algumas indagações: quais seriam as principais diferenças entre o ecossistema informacional antes das redes sociais e atualmente? Ou melhor, quais são as semelhanças e diferenças entre a esfera pública clássica e a esfera pública digital? Vivemos em tempos de polarização dos debates e extremismos ideológicos causados pelas redes sociais ou são apenas ondas políticas passageiras? O que o Estado pode fazer para preservar ou fomentar um ecossistema informacional saudável? Como as empresas que administram redes sociais podem contribuir? Como nós, cidadãos conectados, podemos ajudar? Sob o ponto de vista metodológico, essas indagações podem ser consideradas como as principais perguntas que o trabalho tentará responder, embora não tenhamos expectativa de esgotá-las.

É crucial que se destaque a dificuldade em estudar os temas abordados. Muito embora haja significativo volume de recentes produções acadêmicas nacionais e internacionais sobre a desinformação (o que por si só já se mostrou uma complexidade), e também sobre prováveis vindouras regulações ao direito de liberdade de expressão *online*, para o completo entendimento da questão, é fundamental percorrer outros debates tangentes, muitas vezes carentes de legislação específica e não necessariamente afetos aos usuais campos do Direito. Para somar a essa dificuldade, frisa-se que boa parte das conclusões científicas consideradas aqui está baseada em períodos curtíssimos de observação do comportamento humano nas redes, bem como da interferência das mídias sociais nas esferas públicas democráticas. Portanto, a necessidade de congregar debates tangentes para explicar o fenômeno da desinformação, fenômeno no qual estão inseridas as *fake news* num contexto singular, acaba conferindo um certo grau experimental a todas as iniciativas na área, não sendo diferente para este trabalho. Em suma, a interdisciplinaridade, a

falta de acesso a informações privadas (o problema das *black boxes*) e a novidade do tema representam, num primeiro olhar, as principais dificuldades enfrentadas.

Por este motivo, baseando-se em uma revisão concisa, mas não exaustiva, de trabalhos científicos e não científicos que estudam as consequências da ampliação da esfera pública para o mundo digital, e das recentes posições regulatórias adotadas no cenário global, este ensaio busca contribuir com a construção de uma nova forma de infraestrutura legal e administrativa para a sociedade brasileira - cada vez mais conectada ao mundo digital, e constantemente (re)afetada por inovações tecnológicas.

Assim, considerando a desinformação como um ponto de partida, este trabalho pretende analisar, mais especificamente, questões relacionadas à regulação das redes sociais. Nesta esteira, sabendo que o combate às *fake news* e a outros *online harms* pode atentar contra a liberdade de expressão e ao funcionamento orgânico de uma esfera pública democrática, espera-se traçar um panorama geral para pensarmos um novo sistema de salvaguardas para o ecossistema informacional, especialmente importante em períodos eleitorais.

Para isso, no primeiro capítulo, analisa-se alguns mecanismos do sistema de captação da atenção praticado nas redes sociais, ou seja, algumas características da Internet que potencializam a disseminação de notícias falsas. Destaca-se, primeiro, que as redes sociais são, além de plataformas de interação, verdadeiros canais de distribuição de conteúdo informacional, e por isso são consideradas as novas intermediárias da comunicação, porém intermediárias com características diferentes, pois capazes de criar um ecossistema de informação nunca antes experimentado. Neste ponto fala-se sobre algumas particularidades das Sociedades Algorítmicas: os filtros-bolha e as cascatas cibernéticas. Em um segundo item, debate-se o conceito de *micro-targeting* e a personalização da rede, levantando-se os principais questionamentos deixados pelo caso Cambridge Analytica. Em um terceiro momento, debate-se como os modelos de negócios-online, baseados em cliques, contribuem para o problema, e como a indústria da atenção e os atores políticos mal intencionados podem usar robôs para manipular o debate público. Aqui, são feitos breves apontamentos sobre por que a personalização da rede e o uso desses mecanismos de captação de atenção têm causado o esvaziamento do ecossistema informacional, aumentando a polarização e o extremismo em sociedades conectadas, o que é extremamente deletério para a democracia. Ao final deste primeiro capítulo, e preparando-me para adentrar nas discussões sobre a esfera pública digital, abre-se espaço para conectar esses temas com a ascensão de um novo modelo de propaganda pelo mundo, suas principais características, e os sólidos indícios de sua

utilização ao redor do mundo e também no Brasil: o modelo russo de guerra cibernética pela desinformação.

No segundo capítulo do trabalho, explica-se a natureza contestatória do conceito de esfera pública para concluir que, embora a concepção clássica de esfera pública de Habermas tenha moldado o conceito, as pesquisas sobre esfera pública digital são as que potencialmente mais podem contribuir com os debates emergentes. Assim, coloca-se brevemente a noção habermasiana de esfera pública burguesa, traçando-se as bases necessárias para discutir as principais diferenças entre a noção clássica e a esfera pública digital. Neste ponto, explica-se que, se antes tínhamos uma burguesia intelectual com privilégio de acesso aos meios de comunicação tradicionais, os quais detinham *standards* editoriais mínimos, hoje, diferentemente, assistimos à ascensão do amadorismo na formação da opinião pública, na medida em que qualquer indivíduo conectado é um potencial formador de opinião, e o faz por meios comunicacionais que não possuem mais (necessariamente) padrões de *gatekeeping*.

No terceiro capítulo, partindo-se dessas comparações sobre as diferentes formas de comunicação entre a esfera pública clássica e a digital, são postos os principais *players* e *stakeholders* do atual ecossistema informacional, os quais serão afetados por regulações. Apresentando, então, os usuários, o Estado-regulador e os Intermediários da Internet, monta-se o panorama regulatório dos novos intermediários da comunicação, diferenciando a “*Old School Speech Regulation*” da “*New School Speech Regulation*”. Destaca-se, aqui, que a hibridez desses novos intermediários dificulta a transposição de regulações já existentes para o mundo digital.

Nesta parte do trabalho, explica-se, em poucas palavras, a ideia de regulação e autorregulação regulada para depois expor-se brevemente o regime de isenção de responsabilidade das plataformas nos Estados Unidos, Brasil e Alemanha, os quais historicamente permitiram o desenvolvimento tecnológico do setor. Em seguida, para adentrar ao debate regulatório, destaca-se alguns comportamentos prejudiciais (*online harms*) praticados nas plataformas, frisando-se que a veiculação de *fake news* é apenas um dentre os vários comportamentos com lacunas legais. Aqui, abrimos tópico específico para comentar sobre a tendência internacional em responder a problemas digitais emergentes com a legislação já existente, minimizando-se, assim, a criação de novos tipos legais, os quais encontram dificuldades impostas pela arquitetura do código.

No quarto capítulo, considerando a premissa exposta nas partes anteriores, qual seja, a de que as formas de interação nas plataformas sociais vêm sufocando a cultura democrática que elas mesmas visam promover justamente pelo desdobramento da própria liberdade de expressão, sem freios e restrições, dentro do

ecossistema informacional na arquitetura de controle da atual Sociedade Algorítmica, passa-se a analisar a *public forum doctrine*. A “doutrina do fórum público” é uma ferramenta analítica utilizada na jurisprudência da Suprema Corte norte-americana para interpretar os limites da Primeira Emenda, a fim de determinar a constitucionalidade das restrições de fala implementadas em propriedades públicas. Escolheu-se abordar a *public forum doctrine* porque qualquer regulação imposta às plataformas representa, em certa medida, restrições à liberdade de expressão, sendo que a evolução histórica da doutrina do fórum público acabou construindo balizamentos regulatórios que, à parte eventuais maus usos, conduziram bem o embate entre a liberdade de expressão e a democracia, pelo menos até aqui, de modo que podemos recorrer a elas para pensarmos regulações vindouras.

Diante das lições norte-americanas, comenta-se brevemente sobre uma decisão judicial que representa um marco nas discussões sobre regulação da liberdade de expressão na esfera pública digital, pois aplica essa mesma doutrina do fórum público, desenvolvida para espaços físicos reais, para proibir um Presidente de bloquear usuários de sua conta no Twitter, ou seja, na esfera digital. Considerando que o argumento utilizado para proibir o bloqueio foi o de que o “espaço interativo” da conta privada do Presidente constitui um fórum público, a decisão abre caminho para a transposição de balizamentos regulatórios do mundo real para o virtual. Ao entender que esse “espaço virtual” é um espaço público, podendo ser atingido pelas doutrinas do “fórum público” estabelecidas pela Suprema Corte norte-americana, segundo as quais restrições com base em conteúdo do discurso político são inconstitucionais sob o manto da Primeira Emenda, a decisão da Corte Federal pode representar um excelente embasamento para pensarmos um modelo de autorregulação regulada que empodere as intermediárias a governarem suas plataformas, ou seja, a esfera pública digital, tendo como base a construção legal e jurisprudencial das restrições possíveis e impossíveis à liberdade de expressão já desenvolvidas.

No quinto capítulo, apresenta-se brevemente a lei alemã para a melhoria da aplicação da lei nas plataformas, a *NetzDG*, destacando-se algumas das principais críticas feitas à lei, em especial o perigo de *overblocking* e a aplicação privada do Direito. Em contrapartida, destaca-se também como a lei acaba combatendo esse perigo por meio de uma difusão de direitos e deveres expressos dos usuários e através da ativação da esfera pública digital. A questão posta é que, se de um lado a regulação direta de uma área tão sensível quanto aos “discursos” carrega severos riscos constitucionais, de outra qualquer intervenção direta das e nas plataformas também levanta preocupações sobre censura estatal ou corporativa. Assim, permitir que empresas ou governos determinem a legalidade dos discursos - online ou no

mundo real - certamente esbarra em questões sobre liberdade de expressão e democracia.

Por fim, faz-se uma síntese objetiva dos principais temas abordados e das ideias defendidas ao longo do trabalho, resumindo as conclusões atingidas, e buscando contribuir com um modelo de política regulatória, provavelmente vindoura no Brasil. A título de curiosidade, e correndo o risco inerente a todas as ousadias, comento sobre debates tangentes que, igualmente, orientarão o futuro da democracia, bem como das democracias nas redes.

O questionamento que parece resumir o problema que enfrentaremos aqui é: como podemos conceber uma regulação das redes sociais que, ao mesmo tempo em que cria um ecossistema de informação que valoriza e promove a verdade, não fira de morte a liberdade de expressão?

## 1. O ESAZIAMENTO DO ECOSSISTEMA INFORMACIONAL: COMO É PRODUZIDA A DESINFORMAÇÃO ?

O fenômeno da desinformação obriga-nos a enfrentar o problema sob a ótica da liberdade de expressão, mas também sob o binômio *fake news* e democracia.

Por essa razão, agrupar os complexos tipos de desinformação que fluem no atual ecossistema informacional pode ser um erro, de forma que vale separar a “*disinformation*” - a criação deliberada de informações falsas, ou seja, quando uma determinada informação nasce intencionalmente falsa - da *misinformation*, que é a veiculação inadvertida de informações falsas ou imprecisas. Enquanto no primeiro grupo as informações têm o objetivo (intenção explícita) de enganar, no outro estão conteúdos que na origem poderiam ou não ser falsos, entretanto foram enquadrados de tal maneira a tornar a informação veiculada altamente enganosa ou ilusória. Em outras palavras, sabendo que o próprio conceito de verdade é relativo, é preciso estar atento para distinguir a verdadeira verdade mal contada, e da mentira.

Portanto, chamar todo e qualquer compartilhamento de mentiras de *fake news* significa ignorar o generalismo da expressão e confundir campanhas sistemáticas de desinformação, potencializadas pela atual arquitetura das redes sociais, com conteúdo enganoso,<sup>1</sup> cuja veiculação sempre esteve presente no mundo.

---

<sup>1</sup> Tendo em vista a dificuldade em definir o que seria uma *fake news*, parece bastante produtivo destrinchar os 7 tipos de desinformação apontados por Claire Wardle, da First Draft - organização sem fins lucrativos que dissemina boas práticas no compartilhamento de conteúdo na internet. São eles: (a) Falsa Conexão - quando manchetes, ilustrações ou legendas não confirmam o conteúdo veiculado; (b) Falso Contexto - quando o conteúdo genuíno é compartilhado com outras informações contextualmente falsas; (c) Manipulação do Contexto - quando a informação ou imagem genuína é manipulada para enganar; (d) Conteúdo Enganoso - uso enganoso de informações para enquadrar uma questão ou indivíduo; (e) Conteúdo Impostor - quando fontes genuínas são imitadas ou transcritas; (f) Conteúdo Fabricado - conteúdo novo, que é 100% falso, criado para ludibriar e prejudicar; (g) Sátira ou paródia - quando não há nenhuma intenção de prejudicar, mas tem potencial para enganar. Disponível em: <<https://medium.com/1st-draft/fake-news-its-complicated-d0f773766c79>> (acessado em 04.02.19).



Essa separação parece ser mais relevante quando se discute as formas de contra-atacar as *fake news*, porque, seja pela via regulatória ou por meio de decisões judiciais que determinem a retirada de conteúdo duvidoso do ar, o combate à *disinformation* parece exigir ações interdisciplinares - até mesmo porque pode ser enxergado como modelo de guerra cibernética -, enquanto tentativas de restringir ideias imprecisas (a *misinformation*) têm maiores chances de esbarrar na liberdade de expressão. Certamente, o cuidado nessa seara não é apenas desejado, mas necessário para o saudável funcionamento da esfera pública democrática.

Apesar de ser um termo novo, as *fake news* evoluíram rapidamente. É de se notar, por exemplo, que durante as eleições a mídia tradicional destacou setores específicos apenas para lidar com a profusão de notícias falsas, fato constantemente abordado nos noticiários.<sup>2</sup> O termo deve ser, então, reservado a hipóteses onde haja a apresentação deliberada de declarações falsas ou enganosas como notícias, ou seja, conteúdo com pretensão verniz jornalístico, consideradas enganosas por seu *design*.<sup>3</sup> O trecho "por seu *design*", aqui, se refere às características sistêmicas do *design* das fontes e canais pelos quais as notícias falsas se propagam na atualidade, as redes sociais, e, desse modo, manipulam os processos cognitivos dos consumidores de informação. Chamaremos essas características sistêmicas de "arquitetura de controle das redes sociais." É crucial lembrar, entretanto, que notícias falsas não podem ser confundidas com outras formas de desinformação pública, pois, naturalmente, essa tão recente e controversa expressão acaba sendo usada de várias maneiras (às vezes conflitantes), tornando sua análise conceitual ainda mais difícil. Como também será destacado mais à frente, já é possível perceber o uso das *fake news* até mesmo como tática político-argumentativa para desviar de jornalismo investigativo de qualidade. A estratégia é simples: quando a notícia não agrada, "é *fake news*".

Mapeando o debate, é crucial ter em mente a relação entre o aumento da desinformação e a queda na confiabilidade dos veículos tradicionais de

---

<sup>2</sup> WAACK, William. Fake News: uma visão político-jornalística. in Abboud, Georges. Nery Jr., Nelson. e Campos, Ricardo (Coordenadores). *Fake News e Regulação*. Coleção Direito do Estado e Transformação. São Paulo: Thomson Reuters Brasil, 2018.

<sup>3</sup> GELFERT, Axel. *Fake News: A Definition*. Informal Logic 38 (1) (2018): 84-117.

mídias,<sup>4</sup> como fontes de informação honesta para o consumidor; fato que vem sendo notado ao redor do mundo. No entanto, para não sucumbirmos a uma visão ingênua e que pode vir a parecer nostálgica em relação ao ecossistema informacional do século XX, vale a pena notar que, historicamente, a indústria de jornais e noticiário televisivo sempre foi igualmente acusada de parcialidade, distorção, manipulação e produção de matérias por interesses próprios. Não se tentará argumentar que as mídias tradicionais não propagam inverdades ou não se utilizam do seu poder comunicacional visando interesses econômicos e políticos, próprios e privados. Muito pelo contrário. Pretende-se, aqui, deixar claro que a imparcialidade jornalística é impossível por natureza, mas que construímos, ao longo do século XX, *standards* e padrões editoriais objetivos minimamente confiáveis.

Para avançar ao debate será necessário considerar que as mídias tradicionais possuem padrões ético-jornalísticos mínimos, bem como padrões editoriais e *gatekeeping* objetivos (sem contar as obrigações editoriais legais de autoria e proibição de publicação anônima), que foram capazes de nos trazer, democraticamente, até aqui. Mesmo que contestáveis, será preciso aceitar essas premissas sobre o ecossistema informacional no século XX.

Em suma, embora a falsificação de notícias já exista há muito tempo, e exista em todas as iterações do avanço tecnológico, desde o telégrafo no século XIX aos algoritmos contemporâneos, a arquitetura de controle das plataformas interativas desencadearam novas possibilidades de fraude, fabricação e manipulação do ecossistema informacional. Aqui, ao avaliar quais tipos de notícias falsas podem ter impacto eleitoral, talvez seja útil diferenciar notícias falsas como sátiras, das notícias falsas com fins lucrativos, da propaganda política e reportagens imprudentes.<sup>5</sup>

Sabendo que a circulação de informações falsas pode redundar em má alocação de recursos, desalinhamento de investimentos empresariais, eleições

---

<sup>4</sup> Sobre o tema ver <<http://www.journalism.org/2016/05/26/news-use-across-social-media-platforms-2016/>> (acessado em 04.02.19). Ver, ainda: David M. J. Lazer, Matthew A. Baum, Yochai Benkler, Adam J. Berinsky, Kelly M. Greenhill, Filippo Menczer, Miriam J. Metzger, Brendan Nyhan, Gordon Pennycook, David Rothschild, Michael Schudson, Steven A. Sloman, Cass R. Sunstein, Emily A. Thorson, Duncan J. Watts & Jonathan L. Zittrain, The Science of Fake News, *Science*, 9, Mar, 2018.

<sup>5</sup> PERSILY, Nathaniel. *Can Democracy Survive the Internet?*, 28 *Journal of Democracy* 63 (2017).

mal informadas, e uma série de outros danos que vêm sendo notados ao redor do mundo, sob o ponto de vista democrático o problema concentra-se no fato de o atual ecossistema informacional operar de maneira disfuncional para a esfera pública democrática. E isso vem culminando na incapacidade dos indivíduos de se envolverem em discursos racionais e bem informados sobre a sociedade, na polarização e no extremismo ideológico.

Tendo esse cenário em vista, neste capítulo serão introduzidos debates-chaves para o completo entendimento da disruptiva mudança no ecossistema informacional ocorrida a partir da profusão das redes sociais - que também serão chamadas de plataformas interativas, ou intermediárias da Internet.

Infelizmente, é necessário ressaltar que, embora a quantidade de notícias falsas esteja aumentando,<sup>6</sup> a compreensão científica de como e por que elas se espalham, baseia-se, atualmente, em análises pontuais, em vez de serem em larga escala e multidisciplinares. Portanto, os primeiros desafios estão ligados a uma maior colaboração científica a partir das perspectivas das empresas de tecnologia e da academia, havendo, inclusive quem argumente uma responsabilidade ética e social, transcendendo as forças do mercado, para que as plataformas contribuam com os dados (que possam contribuir e não estejam totalmente protegidos por leis de propriedade industrial) para pesquisas sobre a veiculação de informações falsas, especialmente sobre comportamento e psicologia humana.

Entender como as notícias falsas se espalham é o primeiro passo para contê-las.

---

<sup>6</sup> SOROUGH, Vosoughi. ROY, Deb. ARAL, Sinan. *The spread of true and false news online*. Science, v. 359, 9, Mar. 2018, pp.1146-1151. Disponível em: <<http://science.sciencemag.org/content/359/6380/1146.full>> (acessado em 04.02.19). Tradução livre do autor: "A informação falsa difunde-se significativamente mais, mais depressa, mais profundamente e mais amplamente do que a verdade em todas as categorias de informação, e os efeitos foram mais intensos para notícias políticas falsas do que para falsas notícias sobre terrorismo, desastres naturais, ciência, lendas urbanas ou informações financeiras. Descobrimos que as notícias falsas eram mais inovadoras do que as notícias verdadeiras, o que sugere que as pessoas eram mais propensas a compartilhar informações novas."

## 1.1. Pressupostos teóricos: sobre a arquitetura de controle das redes sociais

Há cerca de 20 anos, as pessoas ainda usavam *modems* e acessavam a Internet por seus telefones residenciais. Essa realidade alterou-se tanto que, muito embora seja difícil precisar o marco temporal da mudança, atualmente podemos considerar que vivemos, segundo Jack Balkin, professor de Direito Constitucional e Liberdade de Expressão da Faculdade de Direito de Yale, em uma Sociedade Algorítmica.<sup>7</sup>

O que quero dizer com a Sociedade Algorítmica? Quero dizer, uma sociedade organizada em torno da tomada de decisão social e econômica por algoritmos, robôs e agentes de inteligência artificial; que não só tomam as decisões, mas também, em alguns casos, as realizam.<sup>8</sup>

É bom guardar esse conceito de “Sociedades Algorítmicas”, pois o termo será constantemente trazido ao longo do trabalho para remeter ao fato de que muitas vezes não fazemos qualquer escolha quanto ao conteúdo informacional que consumimos, e isso deve ser severamente considerado para pensarmos modelos de salvaguardas contra a desinformação.

De qualquer maneira, já se sabe que os algoritmos de busca são fundamentais para nossa experiência na internet, pois, sem eles, teríamos que, pelo menos no que tange ao consumo de informação, escolher dentre os milhões de comentários e notícias postadas diariamente, quais de fato nos interessam.<sup>9</sup> Sem dúvida, esse trabalho seria hercúleo e praticamente

<sup>7</sup> Sobre o tema ver: BALKIN, Jack M. *Free Speech in the Algorithmic Society: Bid Data, Private Governance, and New School Speech Regulation*. Disponível em: <[https://lawreview.law.ucdavis.edu/issues/51/3/Essays/51-3\\_Balkin.pdf](https://lawreview.law.ucdavis.edu/issues/51/3/Essays/51-3_Balkin.pdf)> (acessado em 04.02.19).

<sup>8</sup> Tradução livre do autor do trecho “*What do I mean by the Algorithmic Society? I mean a society organized around social and economic decision making by algorithms, robots, and AI agents; who not only make the decisions but also, in some cases, carry them out.*” in BALKIN, Jack M., *The Three Laws of Robotics in the Age of Big Data* (August 27, 2017). Ohio State Law Journal, Vol. 78, (2017), Yale Law School, Public Law Research Paper No. 592. Disponível em: <<https://ssrn.com/abstract=2890965>> (acessado em 04.02.19).

<sup>9</sup> SUNSTEIN, Cass R. *#Republic: Divided Democracy in the Age of Social Media*. Princeton Univ. Press. 2017. p. 15: “*To some people, unlimited filtering may seem quite strange — a potential product of recent technologies and perhaps even the stuff of science fictio (come true, as it frequently does). But in many ways, it is continuous with what has come before. Filtering is inevitable, a fact of life. It is as old as humanity itself. It is built into our minds. No human being can see, hear, or read everything. In the course of any hour, let alone any day, every one of us*

inviabilizaria o consumo instantâneo de informação como é feito atualmente. Nesse sentido, Cass Sunstein, professor de Direito de Harvard, destaca que em sociedades democráticas, as pessoas são livres para escolherem, por meio de processos de filtragem, com que tipo de conteúdo desejam ter contato.

Com relação ao mundo das comunicações, uma sociedade livre dá às pessoas um grande poder para filtrar materiais indesejados. Somente tiranias forçam as pessoas a ler ou assistir. Nas nações livres, quem lê jornais não lê o mesmo jornal; muitas pessoas não lêem nenhum jornal. Todos os dias, as pessoas fazem escolhas entre revistas com base em seus gostos e ponto de vista.<sup>10</sup>

O simples fato de o cérebro humano tomar essa ou aquela decisão, seja sobre o tipo de comida que escolhermos ingerir, ou o veículo de notícias que escolhermos prestar atenção, estamos realizando algum tipo de filtragem.

É bom ter em mente que os algoritmos - no sentido literal da palavra - nada mais são do que sequências finitas de ações executáveis que visam a obter uma solução para determinado tipo de problema, ou seja, seu uso pode ser entendido como a aplicação de uma fórmula matemática a um determinado contexto, para solucionar um problema. Não obstante, acerca dos algoritmos que nos interessam para estudar o ecossistema informacional, o que ocorre é a aplicação dessas sequências de ações executáveis com a finalidade de filtrar e direcionar conteúdo informacional.

As redes sociais, bem como a maioria dos serviços de comunicação *online* contam com algoritmos para determinar como as notícias - na verdade, qualquer conteúdo - serão disseminadas e consumidas. Ou seja, além da filtragem automaticamente feita na história humana pelo nosso cérebro - que precisa escolher no que prestar atenção -, soma-se a ela a filtragem algorítmica das novas tecnologias.

Portanto, as informações fornecidas no *feed* de notícias do Facebook, na pesquisa do Google e nos tópicos do Twitter são selecionadas e priorizadas

---

*engages in massive filtering, simply in order to make life manageable and coherent. Attention is a scarce commodity, and people manage their own attention in order to ensure that they are not overwhelmed."*

<sup>10</sup> Idem, *Ibidem*. p 16. Tradução livre do autor para o trecho "*With respect to the world of communications, a free society gives people a great deal of power to filter out unwanted materials. Only tyrannies force people to read or watch. In free nations, those who read newspapers do not read the same newspaper; many people do not read any newspaper at all. Every day, people make choices among magazines based on their tastes and point of view."*

por algoritmos altamente complexos, os quais constituem propriedade industrial dessas empresas, e que foram codificados para classificar, filtrar e fornecer conteúdo de uma maneira projetada para maximizar o engajamento dos usuários com o conteúdo e tempo gasto na plataforma. Assim, os algoritmos organizam o grande volume de informações produzidas e compartilhadas nas plataformas, personalizando o conteúdo em torno dos interesses dos usuários e em torno dos interesses de pessoas que têm hábitos e perfis *online* semelhantes.

O que acontece de diferente, no entanto, é que já há algum tempo indivíduos e empresas tentam aplicar esses sistemas de filtragem para fins de marketing e consumo. Por exemplo, todo o modelo de negócios dos buscadores como Google e Bing é construído com base no aumento da classificação de uma empresa nos resultados de pesquisa *online*. Ou seja: se você estiver fora do Brasil e pesquisar as palavras “Rio de Janeiro” + “praia”, muito provavelmente obterá sites relacionados à praia de Copacabana ou Ipanema como resultados. Ou, em exemplo mais corriqueiro e avançado do uso dos algoritmos para consumo: se um usuário acabou de adquirir uma passagem aérea para, digamos, Paris, é economicamente interessante priorizar propagandas de hotéis e hospedagens em suas redes sociais.

Mas, o que é ainda mais novo, e o que será objeto de estudo nesse primeiro capítulo, é a utilização dessas técnicas de negócios e marketing aplicadas à política e à manipulação dos debates na esfera pública dentro de um contexto de incessante busca pela atenção humana. Em outras palavras, é a utilização da desinformação, em um contexto de disputa pela atenção, como arma de controle político e social.

Sobre isto, Tim Wu, professor de Direito da Universidade de Columbia, explica que vários desenvolvimentos tecnológicos e econômicos importantes nas últimas duas décadas transformaram a relativa escassez de atenção da fala e do ouvinte; posições interlocutórias que em tempos de redes sociais acabam se misturando. O primeiro desenvolvimento está associado à popularização da Internet: a queda maciça desde os anos 90 nos custos de ser um comunicador *online*. Este fato também ficou conhecido como "discurso barato" (*cheap speech*), ou "inundação de informação" (*information flood*). Via blogs, micro-blogs ou plataformas interativas como o Twitter ou o Facebook,

praticamente qualquer pessoa, potencialmente, pode disseminar ideias na esfera pública digital. Isso teve várias implicações importantes, que serão vistas no segundo capítulo deste trabalho, como, por exemplo, a ascensão do amadorismo nos debates políticos. O fato é que isso dá às principais redes sociais - que não se consideram parte da imprensa - um papel extremamente importante na construção do discurso público.

Portanto, ao contrário do ecossistema informacional do século XX, vivemos em uma época em que não é mais possível considerar a fala ou a informação como escassas. Na verdade, o que é escasso nas Sociedades Algorítmicas é a atenção dos cidadãos, usuários das redes sociais, que ou não recebem essa ou aquela informação porque seus algoritmos de pesquisa decidiram que elas não eram relevantes para esse indivíduo, ou simplesmente porque sua atenção já foi capturada por outros serviços *online*. Assim, do ponto de vista do ecossistema informacional e em comparação ao século XX, hoje em dia a informação é abundante e falar é fácil; difícil mesmo é ser ouvido. Como consequência disso, o tempo e a atenção dos indivíduos se tornaram *commodities* altamente valorizadas.<sup>11</sup>

O segundo desenvolvimento apontado por Tim Wu, em mais longo prazo, tem sido justamente o surgimento de uma “indústria da atenção” - isto é, “*um conjunto de atores cujo modelo de negócio é a revenda da atenção humana.*”<sup>12</sup> Se, tradicionalmente, esses eram os modelos de negócio das emissoras e jornais, atualmente as principais plataformas interativas da Internet se juntaram a elas, pois, igualmente, procuram maximizar a quantidade de tempo e atenção que as pessoas dedicam a elas. A conclusão é que a ascensão e centralidade da publicidade e propaganda nesses modelos de negócios da indústria da atenção têm o amplo efeito de tornar a atenção do ouvinte cada vez mais valiosa.

O terceiro desenvolvimento apontado pelo pesquisador é o surgimento dos “filtros bolha”,<sup>13</sup> os quais, considerando a importância dada por outros

---

<sup>11</sup> WU, Tim. *Is the First Amendment Obsolete?* Columbia Public Law Research Paper No. 14-573. (2017). p. 6. Disponível em: <<https://ssrn.com/abstract=3096337>>. (acessado em 04.02.19).

<sup>12</sup> Idem, *Ibidem*. p. 7.

<sup>13</sup> Id. p. 8.

pesquisadores às cascatas cibernéticas, como Cass Sunstein,<sup>14</sup> serão analisados conjuntamente. Por enquanto, basta ter em mente que os negócios exploradores de atenção oferecem ao público um pacote de informações altamente adaptado e filtrado, projetado para atender aos interesses preexistentes do cliente, fornecendo conteúdo que maximiza o “engajamento”, ou seja, informações adaptadas aos interesses de cada usuário. Embora isso pareça relativamente inócuo (dar aos usuários o que eles querem), essa estrutura tem o efeito secundário de exercer um forte controle sobre ao que o ouvinte está exposto e bloquear o conteúdo com que seria improvável de ele se envolver. Portanto, a curadoria de conteúdo algorítmico tem consequências importantes sobre como os indivíduos encontram notícias e outras informações políticas importantes, necessárias para uma democracia saudável e funcional.

Não se pretende aqui pular para o lado dos pessimistas e argumentar que as redes sociais vão implodir a cultura democrática. Pretende-se, na verdade, ajudar a fornecer uma lente através da qual podemos entender, um pouco melhor do que antes, o que torna um sistema de liberdade de expressão bem-sucedido e, sob o contexto da desinformação, quais são as exigências necessárias para termos uma democracia funcional. Essas compreensões nos ajudarão a entender as próprias aspirações de uma nação livre e, assim, melhor avaliar mudanças contínuas no ecossistema de comunicação.

Busca-se indicar o caminho para uma compreensão mais clara da natureza da cidadania e seus pré-requisitos culturais. Ora, seria muito simples dizer que qualquer sistema de comunicação é desejável se e porque permite aos indivíduos verem e ouvirem o que escolhem - que é o que está acontecendo com a personalização da rede. Por óbvio, as opções pela personalização do conteúdo também são boas, e a existência de nichos tem enormes vantagens, como a especialização em temas. No entanto, exposições inesperadas (não escolhidas) e experiências compartilhadas também são fundamentais para qualquer sociedade, mostrando-se urgentes para o Brasil polarizado.

Nesse sentido, para as terras tupiniquins - porque até agora ainda não discordamos sobre a existência dos índios, embora haja quem tente instalar

---

<sup>14</sup> SUNSTEIN, Cass R. *#Republic: Divided Democracy in the Age of Social Media*. Op.cit. Nota 10. Capítulo 4.



controvérsias sobre seu legado e seus direitos - fica o já famoso apelo por “serendipidade”, do Professor de Harvard, Cass Sunstein, de acordo com o qual um sistema de liberdade de expressão funcional deve satisfazer dois requisitos distintos.

Em primeiro lugar, as pessoas devem estar expostas a materiais e conteúdos que não teriam escolhido antecipadamente. Os chamados encontros casuais (ou não planejados) frequentemente envolvem tópicos e pontos de vista que as pessoas não procuram por si mesmas e talvez achem até bastante irritantes - mas que têm o condão de mudar suas vidas. Afinal, quem nunca passou por uma situação totalmente inesperada que mudou completamente os rumos da vida? Pessoas pelas quais nos apaixonamos, e que afetam drasticamente nossos rumos, normalmente entram em nossas vidas por puro acaso, e o acaso é consideravelmente diminuído em sociedade com interações previamente “calculadas”. Como veremos, exposições inesperadas são importantes para combater a fragmentação, a polarização e o extremismo, resultados previsíveis de qualquer situação em que grupos homogêneos trocam apenas com eles mesmos.

Em segundo lugar, para Sunstein, as pessoas devem ter uma ampla gama de experiências comuns. Sem experiências compartilhadas, uma sociedade heterogênea terá muito mais dificuldade em abordar problemas sociais complexos. Elas podem até achar difícil se entender, mas experiências comuns, incluindo enfaticamente as experiências possibilitadas pelas mídias sociais, fornecem uma espécie de “cola social”. Lembram da última Copa do Mundo? Pois bem. Um sistema de comunicação que reduz radicalmente o número de experiências desse tipo criará uma série de problemas, sobretudo devido ao aumento da fragmentação social.

Como pré-condições para o bom funcionamento da democracia, esses requisitos - encontros casuais e experiências compartilhadas - devem ocorrer em qualquer país. Eles são especialmente importantes em uma nação heterogênea - que enfrenta um perigo ocasional de fragmentação, como o caso brasileiro. Eles têm ainda mais importância à medida que muitos países se tornam cada vez mais conectados uns aos outros, e cada cidadão, em maior ou menor grau, se torna um “cidadão do mundo”.

Essa ideia parece por si só controversa, porém, se considerarmos, por exemplo, assuntos como imigração, mudanças climáticas e doenças infecciosas, no cenário internacional, e no cenário político nacional, temas como a reforma da previdência, a flexibilização das leis trabalhistas e o avanço do agronegócio sobre terras indígenas, em que raramente as pessoas chegam a acordos, caso nos mantenhamos em *echo chambers* de nossas próprias visões de mundo, será impossível obter uma perspectiva sensata sobre esses assuntos. Até mesmo os problemas mais mundanos, como as interpretações e narrativas sobre fotos e postagens, podem inundar o espaço cibernético, instaurando um clima de discórdia e tumulto informacional.

Espera-se que o leitor - se já não tiver desaparecido para verificar sua página no Facebook ou Instagram - se debruce nas próximas páginas com algumas premissas já estabelecidas. A primeira delas é que vivemos em uma Sociedade Algorítmica na qual os algoritmos de filtragem decidem o conteúdo informacional que chegará até nós. A segunda é que os modelos de negócios praticados por essas plataformas interativas buscam capturar, maximizar e explorar nossa atenção, de maneira que a atenção se tornou uma grande *commodity*. A terceira tem a ver com o fato de que, ao contrário do ecossistema informacional do século XX, não se pode mais considerar a informação e a fala como escassas, muito pelo contrário: diante do barateamento da comunicação, o que vemos é um *overload* informacional, no qual difícil mesmo é separar o conteúdo de qualidade do discurso ruim ou enganoso.

A quarta premissa, a qual tenta-se aprofundar nos subtópicos a seguir, é a ideia de que é possível manusear essa arquitetura de controle das redes sociais para produzir desinformação, e com ela e por meio dela, manipular o debate político nas democracias conectadas.

### 1.1.1 As redes sociais como canais de distribuição de conteúdo informacional: filtros-bolha e cascatas cibernéticas: polarização e extremismo

Desde que o Google introduziu a possibilidade de pesquisa personalizada em 2004, todas as buscas que realizamos no buscador estão associadas às nossas navegações passadas. Isso significa dizer que, quando um usuário realiza uma pesquisa, os resultados não se baseiam apenas na relevância de cada página na web, mas também em uma série de outros sinais de nossa “personalidade” que deixamos espalhados pela rede. De um lado, isso proporciona uma experiência personalizada que tende a aumentar a relevância e a precisão dos resultados da pesquisa para um usuário específico, mas, por outro, essa filtragem também pode ter alguns efeitos colaterais, como criar os “filtros-bolha” e as “cascatas cibernéticas”.

O que se denominou chamar de “filtros-bolha” (*filter bubbles*) pode ser definido, segundo Eduardo Magrani, citando Eli Pariser,<sup>15</sup> como

um conjunto de dados gerado por todos os mecanismos algorítmicos utilizados para se fazer uma edição invisível voltada à customização da navegação on-line. Em outras palavras, é uma espécie de personificação dos conteúdos da rede, feita por determinadas empresas como o Google, através de seus mecanismos de busca, e redes sociais como o Facebook, entre diversas outras plataformas e provedores de conteúdo. Forma-se, a partir das características de navegação de cada pessoa, um universo particular on-line, tanto acessível, quanto (e principalmente) imposto, condicionando sua navegação. Isto se dá a partir do rastreamento de diversos elementos, dentre eles, a localização do usuário e o registro dos cookies, dados de acesso que consistem nas ‘pegadas digitais’ deixadas ao se transitar e se manifestar pelos ambientes on-line.<sup>16</sup>

Portanto, o resultado de cada um de nós estarmos, involuntariamente, inseridos em um filtro-bolha meticulosamente definido a partir de nossas preferências e costumes de navegação faz com que, ao buscarmos por uma

---

<sup>15</sup> Em 2011, Eli Pariser popularizou a teoria dos “filtros-bolha” na qual ele postulou que, devido aos efeitos de uma filtragem algorítmica, os usuários da Internet provavelmente receberão informações que se aproximam de seus interesses existentes e, na verdade, são isolados de diferentes pontos de vista. Em 2018, continuamos a obter evidências sobre o fenômeno (PARISER, Eli. *The Filter Bubble: What the Internet is Hiding from You.*)

<sup>16</sup> MAGRANI, Eduardo. *Democracia conectada: a internet como ferramenta de engajamento político-democrático*. Curitiba: Juruá, 2014, p. 118.

informação na rede, sejamos direcionados a um tipo de conteúdo específico, enquanto outra pessoa qualquer, cujos diferentes hábitos foram igualmente filtrados por seus próprios algoritmos, se depare com um conteúdo bastante diverso.

Como resultado disso, os usuários se distanciam de informações que discordam de seus pontos de vista, isolando-se efetivamente em suas próprias bolhas culturais ou ideológicas, de maneira que os filtros-bolha podem ser, definitivamente, considerados como isoladores intelectuais.<sup>17</sup>

A questão dos filtros-bolha ganha ainda mais importância quando analisada sob a ótica das cascatas de informação, extremamente comuns no ecossistema informacional *online*. Ora, se os algoritmos de controle são capazes de dividir as pessoas em nichos que, por exemplo, curtem esse ou aquele artista, e a partir daí oferecer-lhes artistas ou gêneros musicais semelhantes, também conseguem tratar estes dados com finalidades diversas - como, por exemplo, uma análise acerca das inclinações políticas -, e criar verdadeiros nichos políticos. A forma pela qual isso é feito será explicada no tópico seguinte, porém, para compreendermos as mazelas de se estar inserido em filtros-bolha, é crucial entender como as pessoas tendem a se posicionar diante de ambientes de informação personalizada e nichos políticos balcanizados, ou seja, nichos fragmentados e polarizados.

Cass Sunstein embasado em uma ampla gama de estudos sobre ciência comportamental e behaviorismo *online* e *offline*, argumenta que diante de um público polarizado, após a circulação de informação dentro do próprio grupo, ou na ocorrência de deliberação em grupo não plural, a consequência é um aumento da fragmentação.<sup>18</sup> A consequência disso é o aprofundamento da polarização e do extremismo político.

Imaginemos a hipótese de dois grupos políticos distintos: aqueles que acreditam no aquecimento global e aqueles que acreditam que o aquecimento global é uma farsa, uma invenção de ativistas ambientais, ou até mesmo um simples ciclo pelo qual o planeta Terra passa. Nestes ambientes de público

---

<sup>17</sup> Informação disponível em: <<https://techcrunch.com/2018/06/01/how-instagram-feed-works/>> e <[https://www.huffingtonpost.in/ishtiaque-hossain/filter-bubbles-are-shrinking-our-minds\\_a\\_21469747/](https://www.huffingtonpost.in/ishtiaque-hossain/filter-bubbles-are-shrinking-our-minds_a_21469747/)> (acessados em 04.02.2019).

<sup>18</sup> SUNSTEIN, Cass R. *#Republic: Divided Democracy in the Age of Social Media*. Op.cit. Nota 10. Capítulo 4.

balcanizado, em que os grupos estão projetando suas próprias ideias e pacotes de comunicação preferidos, a consequência não será apenas a manutenção das posições, mas, o que é pior: um severo agravamento da polarização à medida que os membros do grupo se movem para pontos mais extremos em comparação com suas tendências iniciais.

Mesmo que a maioria de nós não se isole por completo de outros pontos de vista, algumas ou muitas pessoas farão e estarão fazendo exatamente isso. Basta pararmos para pensar que um tema recorrente da atualidade são as brigas e exclusões em grupos familiares ou de amigos (*online* ou na vida real), originadas a partir de discussões políticas. Isso é suficiente para ocorrer a polarização e causar sérios riscos sociais. Em geral, são precisamente as pessoas mais propensas a dispensar visões opostas que mais precisam ouvir tais visões. Nessa perspectiva, as redes sociais tornam mais fácil ouvir opiniões de outros que pensam da mesma maneira, e isolando-se das visões concorrentes, o que é ainda mais intensificado no caso de usuários que simplesmente excluem os vínculos *online* com “opositores políticos”. Só por essa razão as plataformas interativas já se mostram um terreno fértil para a polarização e potencialmente perigosas para a democracia e a paz social.

Fazendo uma leitura complementar da obra “#Republic: Divided Democracy in the Age of Social Media” e do artigo “Fragmentation and Cybercascades”, Sunstein explica que há dois principais motivos para a polarização de grupos. O primeiro deles tem a ver com a persuasão, e por isso está mais relacionado às cascatas informativas. Já o segundo está conectado ao que se convencionou chamar de cascatas reputacionais, e leva em conta o fato de que as pessoas querem ser percebidas favoravelmente por outros membros do grupo do qual “fazem parte”.

Em uma cascata informacional, as pessoas deixam de confiar em determinado ponto em suas informações ou opiniões privadas. Elas decidem, em vez disso, com base nos sinais transmitidos por outros. Segue-se que o comportamento das primeiras poucas pessoas, ou mesmo uma, pode, em teoria, produzir um comportamento similar de inúmeros seguidores. (...) Podemos também imaginar a possibilidade de cascatas reputacionais, paralelas aos seus irmãos informacionais. Em uma cascata de reputação, as pessoas pensam que sabem o que é certo ou o que é provável que seja certo, mas, não obstante, acompanham a multidão para manter a boa opinião de outros. Mesmo

as pessoas mais confiantes às vezes são vítimas dessa pressão, silenciando-se no processo. Temendo a ira dos outros, as pessoas não podem contestar publicamente práticas e valores que privadamente abominam.<sup>19</sup>

A primeira explicação baseia-se em uma intuição simples: a posição de qualquer indivíduo em qualquer questão é uma função (ou uma média), pelo menos em parte, de quais argumentos parecem convincentes sobre determinado assunto. Considerando então que a posição se move a partir da discussão em grupo (ou do consumo de informação), é provável que ela se mova na direção da posição mais persuasiva defendida dentro do grupo, tomado como um todo. Agora, se os membros do grupo já estiverem inclinados numa certa direção - por exemplo, digamos que não acreditem em mudanças climáticas e aquecimento global -, eles oferecerão um número desproporcionalmente grande de argumentos na mesma direção, e um número desproporcionalmente pequeno de argumentos indo na direção contrária; isso sem contar a qualidade dos argumentos e a ênfase posta. Como resultado, a consequência da discussão será levar as pessoas ainda mais ao extremo da direção de suas inclinações iniciais. Em suma, o fator central por trás da polarização de grupo é a existência de um conjunto limitado de argumentos - já viesados em uma direção particular.

A segunda explicação, envolvendo a comparação social, começa com a sugestão razoável de que as pessoas querem ser percebidas favoravelmente por outros membros do grupo e também se percebem favoravelmente, de maneira que elas frequentemente ajustam suas posições na direção da posição dominante. Suponhamos, por exemplo, que pessoas em um determinado nicho acreditem no aquecimento global e nas mudanças climáticas, e que parecem querer se opor fortemente a tudo o que é contra isso. Se elas estão em um grupo cujos membros também se opõem fortemente a essas coisas, elas podem mudar na direção de uma oposição ainda mais forte depois de ver o que os outros membros do grupo pensam. Sunstein se remete a inúmeros estudos nos quais exatamente esse padrão foi observado. Por óbvio, as

---

<sup>19</sup> Tradução livre do autor do trecho: “*In an informational cascade, people cease relying at a certain point on their private information or opinions. They decide instead on the basis of the signals conveyed by others. It follows that the behavior of the first few people, or even one, can in theory produce similar behavior from countless followers*” in Idem, *Ibidem*. Capítulo 4.

peças não mudam se tiverem uma noção clara do que pensam e não são movidas pelas opiniões dos outros. Porém, a maioria das pessoas, na maioria das vezes, não é tão fixa em seus pontos de vista, ainda mais sobre assuntos complexos, e ultimamente temos visto até nichos que acreditam que o Planeta Terra seria plano, questão que não deveria mais estar sob discussão.

Nesse sentido, a situação preocupa ainda mais se considerarmos que pesquisas recentes indicam que a maioria das pessoas recebe suas notícias por meio das redes sociais: no caso do Brasil<sup>20</sup> cerca de 70%, e no dos Estados Unidos 65%, na combinação de pesquisas.<sup>21</sup>

No artigo *“Fake News e os Caminhos para Fora da Bolha”*, Sérgio Branco, nos alertou de que a personificação dos conteúdos *“nos [dá] aquilo que nós mais gostamos: nós mesmos. E é muito difícil vencer essa tentação narcisista.”* Consequentemente, para o pesquisador, os filtros-bolha acabam limitando a diversidade.<sup>22</sup>

A preocupação aqui é de que, para uma democracia funcionar bem, as pessoas não podem viver em casulos de informação. Elas deveriam estar em contato com uma ampla gama de pontos de vista e ideias. Na verdade, elas deveriam fazer isso mesmo sem querer ver ou ouvir sobre determinados tópicos, pois é isso que faz com que os cidadãos estejam em contato com a pluralidade e as muitas outras visões de mundo. Isso levanta sérias questões sobre nosso comportamento nas redes e os múltiplos usos das redes sociais.

Tendo em vista, portanto, que boa parte da informação que consumimos nos é trazida por meio das redes sociais, e que esses novos transmissores de informação estão personalizando o conteúdo que recebemos,

---

<sup>20</sup> Informação disponível em: <<http://observatoriodaimprensa.com.br/e-noticias/cerca-de-70-dos-brasileiros-se-informam-pelo-facebook/>> (acessado em 04.02.19).

<sup>21</sup> Informação disponível em: <<http://www.journalism.org/2016/05/26/news-use-across-social-media-platforms-2016/>> (acessado em 04.02.19).

<sup>22</sup> *“A fim de satisfazer suas próprias necessidades, inúmeros serão os usuários que optarão por deixar de seguir (dar unfollow) em quem quer que discorde deles, que não curta, nem compartilhe suas publicações ou que simplesmente lhe pareça desinteressante. No limite, um usuário pode ser até mesmo bloqueado. Em contrapartida, poderá (porque o Facebook também o permite) eleger um certo número de amigos cujas postagens serão vistas prioritariamente. Em regra, serão conteúdos com os quais o usuário se identifica e por cujas atualizações anseia.”* in<<https://itsrio.org/wp-content/uploads/2017/08/sergio-fakenews.pdf>> (acessado em 04.02.19).

é crucial entender melhor como esse processo ocorre e a posição em que se encontram as principais plataformas interativas.<sup>23</sup>

Pois bem. A personalização do conteúdo que buscamos é possível a partir da coleta de dados e outras informações que deixamos espalhadas pela rede e seu tratamento realizado por meio do uso de algoritmos.<sup>24</sup> Isso vem sendo feito não apenas pelos buscadores, como Google e Bing, mas de forma bastante similar pelas redes sociais, como o Facebook<sup>25</sup>, o Instagram<sup>26</sup> e o Twitter<sup>27</sup>, em seus “feeds de notícias”. O professor de Harvard, Cass Sunstein, explica que:

O próprio Facebook faz alguma curadoria e o Google também. Vivemos na era do algoritmo e o algoritmo sabe muito. Com o surgimento da inteligência artificial, os algoritmos tendem a melhorar imensamente. Eles aprenderão muito sobre você e saberão o que você quer e/ou prefere antes de você, e melhor do que você. Eles até mesmo conhecerão suas emoções, novamente antes e melhor do que você, e serão capazes de imitar emoções por conta própria. (...) Se o algoritmo sabe que você gosta de certos tipos de música, ele pode saber, com alta probabilidade, de que tipos de filmes e livros você gosta e de quais candidatos políticos serão atraentes para você. E, se souber quais sites você visita, pode saber quais produtos você provavelmente comprará e o que pensa sobre a mudança climática e a imigração. Um pequeno exemplo: o Facebook provavelmente conhece suas convicções políticas e pode informar outros, incluindo candidatos a cargos públicos, sobre o que sabe. Ele categoriza seus usuários como muito conservadores, conservadores, moderados, liberais e muito liberais. Ele faz isso vendo as páginas que você gosta. Se você gosta de certos candidatos de forma favorável ou desfavorável, a categorização é ainda mais fácil. Inclusive, o Facebook não esconde o que está fazendo. Na página “Preferências

---

<sup>23</sup> Há uma discussão interessante sobre a classificação das redes sociais, que será abordada em momento oportuno no segundo capítulo: seriam elas as novas intermediárias da comunicação? Deveriam ser tratadas como a imprensa ou não?

<sup>24</sup> “*The Algorithmic Society features the collection of vast amounts of data about individuals and facilitates new forms of surveillance, control, discrimination and manipulation, both by governments and by private companies. Call this the problem of Big Data.*” in BALKIN, Jack M. *Free Speech in the Algorithmic Society: Bid Data, Private Governance, and New School Speech Regulation*, op. cit. nota 8. p. 3.

<sup>25</sup> “*As publicações que aparecem primeiro são influenciadas por suas conexões e atividades no Facebook. O número de comentários, curtidas e reações recebidos por uma publicação e o seu tipo (foto, vídeo, atualização de status) também podem torná-la mais propensa a aparecer primeiro no seu Feed de Notícias.*”. Disponível em <<https://www.facebook.com/help/166738576721085>> (acessado em 04.02.19).

<sup>26</sup> Disponível em: <<https://www.independent.co.uk/life-style/gadgets-and-tech/news/instagram-feeds-to-go-out-of-order-showing-pictures-according-to-relevance-rather-than-time-using-a6933926.html>> (acessado em 04.02.19).

<sup>27</sup> “*Os Tweets que você provavelmente gostará mais aparecerão primeiro em sua timeline. Nós os escolhemos com base nas contas com as quais você mais interage, nos Tweets com os quais você se envolve e muito mais.*” Disponível em <<https://help.twitter.com/pt/using-twitter/twitter-timeline>> (acessado em 04.02.19).



de anúncios no Facebook,” você pode pesquisar em "Interesses" e, em "Mais", em "Estilo de vida e cultura" e, finalmente, em "Política dos EUA", e a categorização aparecerá.<sup>28</sup>

Sunstein lembra também que, se de um lado é verdade que as redes sociais sabem muito sobre nós, de outro é importante observar que as mídias sociais vieram ao mundo não como órgãos da imprensa livre, não como serviços de notícias e informações, não como mediadores entre cidadão e Estado e nem sequer como mediadores entre cidadãos, mas como maneiras divertidas de socializar. A função das mídias sociais como canais de transmissão de informações, incluindo-se aqui o noticiário político, simplesmente aconteceu.<sup>29</sup> Talvez seja este o motivo pelo qual as próprias plataformas não se consideram como intermediadoras da comunicação como a imprensa, ou talvez esse papel possa ter sido previsto, uma vez que ocorreu paulatinamente. Mas, certamente, não foi com essa finalidade que as plataformas interativas foram concebidas ou aparentemente planejadas; inclusive porque essas plataformas estão em constante mutação - uma regra fundamental do ambiente tecnológico.

Assim, preparadas para isso ou não, a realidade é que essas plataformas são as novas instituições intermediárias do atual fluxo informacional. As organizações políticas e os meios de comunicação tradicionais não ressurgirão para a era da Internet em nada parecido com suas encarnações anteriores, pois, para serem economicamente sustentáveis,

---

<sup>28</sup> SUNSTEIN, Cass R. *#Republic: Divided Democracy in the Age of Social Media*, op.cit. nota 10. p. 13. Tradução livre do autor do trecho: “*Facebook itself does some curating, and so does Google. We live in the age of the algorithm, and the algorithm knows a lot. With the rise of artificial intelligence, algorithms are bound to improve immeasurably. They will learn a great deal about you, and they will know what you want and or will like, before you do, an better than you do. They will even know your emotions, again before and better than you do, and they will be able to mimic emotions on their own. (...) If the algorithm knows that you like certain kinds of music, it might know, with a high probability, what kinds of movies and books you like, and what political candidates will appeal to you. And if it knows what websites you visit, it might well know what products you're likely to buy, and what you think about climate change and immigration. A small example: Facebook probably knows your political convictions and it can inform others, including candidates for public office, of what it knows. It categorizes its users as very conservative, conservative, moderate, liberal, and very liberal. It does so by seeing what pages you like. If you like certain candidates favorably or unfavorably, categorization is easier still. By the way, facebook doesn't hide what it is doing. On the Ad Preferences page on Facebook, you can look under "Interests," and then under "More", and then under "Lifestyle and culture," and finally under "US Politics," and the categorization will come right up.*”

<sup>29</sup> No mesmo sentido WU, Tim. *Is the First Amendment Obsolete?* Columbia Public Law Research Paper N. 14-573.

também precisam aderir ao novo *modus operandi* do mercado e também estão promovendo conteúdo dentro dessas mídias sociais. Neste ponto, vale a pena antecipar um debate travado no terceiro capítulo, acerca da hibridiz das intermediárias da comunicação, ou seja, hoje em dia elas acabam desempenhando vários papéis distintos, o que dificulta o tratamento jurídico da matéria.

No entanto, diferentemente das instituições anteriores, essas novas “empresas de informação” não foram criadas para servir a valores democráticos e, a princípio, não possuem a promoção de um eleitorado bem informado e de mentalidade cidadã como sua finalidade primeira - “mentalidade cidadã” essa que, para as finalidades do debate, pode ser resumida como um agir político, racional, e razoavelmente bem justificado na esfera pública (algo nos termos da participação democrático-deliberativa).<sup>30</sup> Elas foram criadas como plataformas de interação *online*.

Nesse sentido, as redes sociais e outras empresas de tecnologia estão em uma posição estranha: elas argumentam que sua contribuição para a disseminação de *fakes news*, por exemplo, é puramente casual, uma combinação infeliz de traços humanos e plataformas neutras de compartilhamento de informações. No entanto, elas estão simultaneamente argumentando que seus algoritmos e processos empresariais prezam pelos

---

<sup>30</sup> Antecipando um debate do segundo capítulo, cumpre resumir que os principais pensadores sobre o modelo deliberativo de democracia, como Thomas Christiano, Joshua Cohen, Seyla Benhabib e Bernard Manin, argumentam, em conjunto, que a deliberação pública aperfeiçoa a qualidade das decisões na democracia, possuindo um valor intrínseco, e alcançando resultados objetivos. Resumidamente: a deliberação permite aos indivíduos compreenderem melhor os interesses existentes na sociedade, e, por conseguinte, a própria sociedade. Mesmo quando a deliberação pública produz mais desacordos do que acordos, do ponto de vista democrático-institucional isso também pode ser positivo. Quanto aos valores intrínsecos da deliberação, os autores frisam que num ambiente onde as partes deliberam umas com as outras, manifesta-se algum tipo de respeito entre os cidadãos. Somado a isso, o ambiente da deliberação pública melhora as condições cognitivas, habilitando o cidadão a entender melhor o que está em jogo numa decisão política. Nesse contexto, a deliberação mostra-se valiosa instrumentalmente, pois torna-se um instrumento/procedimento para tomar decisões mais bem informadas, pois, considerar uma maior multiplicidade de pontos de vista, em última análise, torna a decisão mais qualificada, e legitimamente mais justificada; pois, plural e procedimentalista. Sobre o tema, ver: COHEN, Joshua. *Deliberation and democratic legitimacy*. In: hamlin, Alan; Pettit, Philip (Eds.). *The Good Polity*. Oxford: Blackwell, 1989, p. 17-34. Ver também artigos de CHRISTIANO, Thomas. “*A importância da deliberação pública*” e “*Procedimento e substância na democracia deliberativa*”, bem como BENHABIB, Seyla. “*Sobre um modelo deliberativo de legitimidade democrática*”, os três In: MELO, Rúrion Soares; WERLE, Denílson Luiz (Eds.). *Democracia Deliberativa*. São Paulo: Editora Singular, 2007.

direitos de liberdade de expressão em primeiro plano, sendo suas plataformas ambientes de livre compartilhamento de ideias.

Então, pode-se considerar que as missões do jornalismo e das plataformas estão, de certa forma, em desacordo: enquanto o jornalismo tenta fornecer os fatos necessários, mas às vezes desconfortáveis, bem como o contexto pertinente para dar sentido a eles e manter o cidadão bem informado, as redes sociais visam fornecer informações desejadas e divertidas para manter seus usuários felizes e interessados, captando o máximo de atenção, pelo maior tempo possível.

Não está claro, portanto, se existe uma maneira de vincular essas missões aparentemente separadas, e se os novos distribuidores de conteúdo estão preparados para assumir o papel de provedores de informação do século XXI. Nasce então um questionamento importante: como compatibilizar os interesses jornalísticos, e os interesses das plataformas digitais se, pelo menos à primeira vista, eles são conflitantes?

Em suma, essa dificuldade de compatibilização e a existência dos filtros-bolha e das cascatas cibernéticas nos deixam reféns dos algoritmos desenvolvidos pelas mídias sociais, criando obstáculos sérios ao acesso à informação.

Nessa perspectiva, adentra-se o tema da filtragem algorítmica e o alto grau de personalização de conteúdo, cabendo reportar como o recente caso Cambridge Analytica transportou a discussão para a ótica eleitoral.

### 1.1.2 Micro-targeting e o caso Cambridge Analytica

É sob o enfoque da personalização da rede e do direcionamento de informação para nichos políticos que o caso Cambridge Analytica chocou o mundo, pois, por meio do tratamento dos dados pessoais de usuários coletados no Facebook, a empresa conseguiu desenvolver perfis “psicográficos” para cada indivíduo, e permitir um direcionamento de discurso ainda mais subjetivo.

O mapeamento de preferências e costumes na rede é uma estratégia que já vem sendo utilizada em larga escala pelo marketing digital e pelo comércio *online*. Individualizando-se ao máximo os consumidores, as plataformas são capazes de definir nossas preferências de consumo e “empurrar” (*nudge*) produtos e serviços, tudo isso através de um lucrativo mercado de comércio de dados.

Em 2013, Michal Kosinski, David Stillwell e Thore Graepel, todos eles pesquisadores da Universidade de Cambridge, no Reino Unido, à época, publicaram um artigo<sup>31</sup> intitulado “Private traits and attributes are predictable from digital records of human behavior” que concluía que as coisas que “curtimos” no Facebook podiam ser usadas para prever traços de nossa personalidade. O estudo se baseou em cinco traços bem estabelecidos da personalidade humana, os chamados “Big Five”: agradabilidade, neuroticismo, abertura a novas experiências, extroversão e conscienciosidade.

Os pesquisadores chegaram a essas conclusões ao introduzir um aplicativo no Facebook que avaliou traços da personalidade (por meio de um teste/quiz) e outras características pessoais dos usuários. Em seguida, os pesquisadores cruzaram esses dados com informações sobre o que esses usuários gostavam, conseguindo traçar um padrão, ou seja, realizar uma microsegmentação em nichos e prever personalidades. Esse modelo foi “copiado” pela empresa Cambridge Analytica – que por sinal já foi fechada - e

---

<sup>31</sup> KOSINSKI, Michal. STILLWELL, David. GRAEPEL, Thore. *Digital records of behavior expose personal traits*. Proceedings of the National Academy of Sciences Apr 2013, 110 (15) 5802-5805. Disponível em: <<http://www.pnas.org/content/110/15/5802#ref-15>> (acessado em 04.02.19).

transportado para fins eleitorais, nas primárias do Partido Republicano pelo oponente de Trump, Ted Cruz, e depois pelo próprio Trump.

Usando esses perfis psicográficos, é possível não apenas identificar quais eleitores são mais propensos a serem balançados por certas causas, candidatos, assuntos e até frases de efeito, mas também manipular a informação para prever e alterar o comportamento futuro desses potenciais eleitores, por meio de propaganda personalizada, adaptativa e viciante. O objetivo não é apenas avaliar corretamente o provável comportamento do eleitor, mas, quando necessário, mudá-lo, explorando suas emoções.

Acerca do aproveitamento das emoções, é importante notar que uma das facetas da desinformação é justamente a manipulação das emoções e dos instintos humanos via conteúdo com pretensão verniz jornalístico, ou seja, através de frases apelativas, cores e designs chamativos. Nesse sentido, estudos mostram, inclusive, que estágios emocionais podem ser transferidos para outros, podendo haver, de fato, contágio emocional pelas redes sociais. Assim, é possível levar um indivíduo a experimentar as mesmas emoções sentidas por aqueles que os rodeiam nas redes.<sup>32</sup>

Embora seja difícil compreender como isso é feito, a partir das aplicações descritas pelos próprios pesquisadores que desenvolveram o método, o qual inicialmente não fora pensado para finalidades eleitorais, pode-se ter uma noção do potencial da descoberta:

A previsão dos atributos e preferências individuais dos usuários pode ser usada para melhorar vários produtos e serviços. Por exemplo, os sistemas e dispositivos digitais (como lojas online ou carros) podem ser projetados para ajustar seu comportamento de modo a se adequar melhor ao perfil inferido de cada usuário. Além disso, a relevância das recomendações de marketing e produtos poderia ser melhorada adicionando-se dimensões psicológicas aos modelos atuais do usuário. Por exemplo, anúncios de seguros online podem enfatizar a segurança ao enfrentar usuários emocionalmente instáveis (neuróticos), mas enfatizar ameaças potenciais ao lidar com usuários emocionalmente estáveis. Além disso, registros digitais de comportamento podem fornecer uma maneira conveniente e confiável de medir traços psicológicos. A avaliação automatizada, baseada em

---

<sup>32</sup> KRAMER, Adam. GUILLORY, Jamie. HANCOCK, Jeffrey. *Experimental Evidence of Massive-Scale Emotional Contagion Through Social Networks*. Proceedings of the National Academy of Sciences of the United States of America. 111. 10.1073/pnas.1320040111. Disponível em [https://www.researchgate.net/publication/262813340\\_Experimental\\_Evidence\\_of\\_Massive-Scale\\_Emotional\\_Contagion\\_Through\\_Social\\_Networks](https://www.researchgate.net/publication/262813340_Experimental_Evidence_of_Massive-Scale_Emotional_Contagion_Through_Social_Networks) (acessado em 04.02.19).

grandes amostras de comportamento, pode não apenas ser mais precisa e menos propensa a fraudes e deturpações, mas também permitir a avaliação ao longo do tempo para detectar tendências. Além disso, a inferência baseada em observações do comportamento registrado digitalmente pode abrir novas portas para a pesquisa em psicologia humana.<sup>33</sup>

Assim, em termos simples, a interação entre quem produz o conteúdo e quem recebe não se parece nem um pouco com cenários antigos de fluxo de informação, e a transposição do *microtargeting* a partir da microsegmentação em nichos para o ambiente eleitoral é preocupante, pois, agora, o emissor entende as personalidades de seus interlocutores e, em seguida, pode usar essas informações para direcionar anúncios e propaganda política às pessoas com uma maior probabilidade de obter resposta positiva, ou até mesmo alterar palavras e expressões ao se comunicar com grupos propensamente menos receptivos a determinada ideia.

---

<sup>33</sup> Tradução livre do autor do trecho “*Predicting users’ individual attributes and preferences can be used to improve numerous products and services. For instance, digital systems and devices (such as online stores or cars) could be designed to adjust their behavior to best fit each user’s inferred profile. Also, the relevance of marketing and product recommendations could be improved by adding psychological dimensions to current user models. For example, online insurance advertisements might emphasize security when facing emotionally unstable (neurotic) users but stress potential threats when dealing with emotionally stable ones. Moreover, digital records of behavior may provide a convenient and reliable way to measure psychological traits. Automated assessment based on large samples of behavior may not only be more accurate and less prone to cheating and misrepresentation but may also permit assessment across time to detect trends. Moreover, inference based on observations of digitally recorded behavior may open new doors for research in human psychology.*” in op. cit. nota 32.

### 1.1.3 O uso de robôs e os modelos de negócio online

De acordo com estimativas<sup>34</sup> do provedor de serviços em nuvem Imperva Incapsula, os *bots* representaram 51,2% de todo o tráfego da web em 2016. Isso significa que, além da atividade humana na Internet e do uso saudável de robôs,<sup>35</sup> atores estatais e não-estatais estão criando cada vez mais contas automatizadas nas mídias sociais para manipular discursos na esfera pública *online*.

Segundo o relatório “Freedom on the Net 2017 - Manipulating Social Media to Undermine Democracy”<sup>36</sup>, da Freedom House, em pelo menos 20 países padrões característicos de atividade online sugeriram o uso coordenado de tais *bots* para influenciar o cenário político. O uso desses robôs na política é feito com três objetivos principais: (1) fingir apoio de base para determinado ator político (tática também conhecida como “*astroturfing*”), (2) desagrupar oponentes e (3) afastar a atenção de assuntos polêmicos, manipulando os “*trend topics*”.

No contexto de disputas político-eleitorais, os robôs podem ser manuseados para distorcer a dimensão de movimentos políticos, manipular e radicalizar debates na esfera pública, e criar falsas percepções sobre disputas e consensos nas redes sociais.

É possível medir o poder de uma *fake news* pela viralidade da mentira que ela propaga, pela velocidade com que é disseminada e, conseqüentemente, pelo número de pessoas que recebem e acreditam na falsidade. Muitos fatores podem impulsionar a popularidade de uma história, mas o ritmo com que as mentiras podem viajar nas plataformas interativas é

---

<sup>34</sup> Informação disponível em: <<https://www.incapsula.com/blog/bot-traffic-report-2016.html>> (acessado em 04.02.19).

<sup>35</sup> Muitos deles realizam tarefas automatizadas para fins comerciais. Por exemplo, os *bots* agora desempenham um papel vital no monitoramento da integridade dos sites, na solicitação de produtos *online* para checagem de sites e no envio de novos conteúdos, por exemplo. Esses “bons bots” são identificáveis e operados por muitas das maiores empresas de tecnologia, incluindo Amazon, Facebook, Google e Microsoft. *Bots* mal-intencionados, no entanto, não são identificáveis por design e compõem a maioria das atividades de *bot* desde 2013.

<sup>36</sup> FREEDOM HOUSE. *Freedom on the Net 2017: Manipulating Social Media to Undermine Democracy* (November 2017). Disponível em: <[https://freedomhouse.org/sites/default/files/FOTN\\_2017\\_Final.pdf](https://freedomhouse.org/sites/default/files/FOTN_2017_Final.pdf)> (acessado em 04.02.19).

certamente potencializado pelo uso coordenado de *bots*, e não apenas por razões diretamente relacionadas ao conteúdo disseminado.

Essas contas automatizadas se apresentam como humanos no Facebook, no Twitter e em outras redes sociais. Como relata Tim Wu, há pesquisas que estimam que o Twitter tenha até 48 milhões de robôs-usuários, e o Facebook já entre 67,65 milhões e 137,76 milhões de usuários falsos, sendo alguns deles aproveitados para ajudar a espalhar desinformação e outras táticas de guerra cibernética, como a perseguição a jornalistas.<sup>37</sup>

No Twitter, por exemplo, contas de *bots* costumam tuitar com frequência, retuitar uns aos outros e disseminar links para conteúdo externo com mais frequência do que as contas operadas por humanos. Por exemplo, a Newsweek noticiou<sup>38</sup> que uma revisão dos seguidores do presidente Donald Trump no Twitter indicou que apenas 51% de seus 30 milhões seguidores eram reais, à época.

No Brasil, recente relatório<sup>39</sup> do InternetLab concluiu que, no Twitter, entre aproximadamente 13 a 47% dos seguidores dos pré-candidatos à Presidência eram *bots*.

Percebe-se, então, que os *political bots* podem moldar a esfera pública *online* para parecer que determinada figura é mais popular do que de fato é ou, ainda, serem utilizados para replicar discursos em série, fazendo parecer que há uma posição majoritária acerca de uma causa quando, na verdade, não há.

Mais especificamente sobre a disseminação de informações falsas, a prática denominada “*astroturfing*” funciona da seguinte forma: múltiplas identidades online e grupos de pressão falsos são usados para induzir o público a acreditar que determinada ideia escolhida representa uma visão comumente aceita, ou seja, uma posição majoritária. Neste cenário, os *bots* são utilizados, por exemplo, para promover falsamente uma notícia com pretensão verniz jornalístico, introduzidas numa variedade de grupos políticos, criando a impressão de amplo apoio popular a uma ideia, política, indivíduo ou

---

<sup>37</sup> WU, Tim. *Is the First Amendment Obsolete?* Columbia Public Law Research Paper No. 14-573, 2017, pp. cit. nota 30. p. 16.

<sup>38</sup> Informação disponível em: <<http://www.newsweek.com/donald-trump-twitter-followers-fake-617873>> (acessado em 04.02.19).

<sup>39</sup> Informação disponível em: <<http://www.internetlab.org.br/wp-content/uploads/2018/07/Relat%C3%B3rio-Bots-ou-n%C3%A3o.pdf>> (acessado em 04.02.19).



produto. Isso é atingido por meio da multiplicação do número de acessos e compartilhamentos potencializados pelos *bots* programados com essas intenções e, paralelamente, por pessoas reais que não se deram ao trabalho de verificar a informação.

Esse uso malicioso de robôs representa uma sofisticada e lucrativa indústria transnacional de "likes" e seguidores, que opera nas sombras da Internet. Nesse sentido, Carlos Affonso Pereira de Souza já nos advertiu que os modelos de negócio dessas plataformas contribuem para o problema.

O atual modelo financeiro das mídias digitais favorece e incentiva o fenômeno das fake news: quanto mais cliques, maior o retorno. A remuneração através de anúncios nem sempre está relacionada ao conteúdo que é disponibilizado, mas sim à quantidade de acessos. Os autores dessas notícias, então, buscam maximizar o tráfego em suas páginas e, conseqüentemente, o lucro. No ano passado, por exemplo, uma pessoa apenas arrecadou trinta mil dólares com o seu site de notícias falsas. A estrutura da Internet tornou a divulgação de fake news economicamente rentável. No passado, não seria possível adquirir tamanha amplitude, pois o custo de distribuição seria enorme, além da dificuldade em construir uma audiência cativa. Por meio das mídias sociais, o cenário se transforma: cada pessoa pode influenciar outras com as postagens de sua própria página.<sup>40</sup>

Assim, o problema dessa estrutura financeira se mostra duplo. A primeira complicação é que as informações falsas (principalmente as que apelam para as emoções) são mais lucrativas do que links com notícias verdadeiras, porque aquelas são mais chamativas (*click baits*). O segundo obstáculo é que, dentro desse mercado da atenção, os algoritmos utilizados pelas plataformas para nos direcionar conteúdo possuem um altíssimo valor de mercado, estando protegidos por leis de propriedade intelectual e por protocolos de segurança empresariais.

Por um lado, as plataformas interativas precisam do dinheiro da publicidade, sendo que nos modelos de negócio praticados por elas (mercado de captação de atenção) o dinheiro é gerado por cliques. Como um estudo do

---

<sup>40</sup> SOUZA, Carlos Affonso. PADRÃO, Vinicius. *Quem lê tanta notícia (falsa)? Entendendo o combate contra as "fake News*. Disponível em: <<https://feed.itsrio.org/quem-l%C3%AA-tanta-not%C3%ADcia-falsa-entendendo-o-combate-contra-as-fake-news-70fa0db05aa5>> (acessado em 04.02.19).

Massachusetts Institute of Technology (MIT) indicou, as *fake news* conseguem mais cliques e ganham mais dinheiro do que as notícias verdadeiras.<sup>41</sup>

Por outro lado, a publicidade por si só não é o único problema. A imprensa tradicional também precisa de verba de propaganda e não está no caminho de virar intermediária veiculadora de *fake news* (ou ainda não, pelo menos). Assim, o segundo problema está mais relacionado à falta de transparência na maneira como os algoritmos trabalham para canalizar informações e notícias através dos *feeds* das redes sociais ou dos provedores de busca, bem como a nossa impossibilidade de questionar ou alterar o tratamento desses dados, uma vez que os algoritmos acabam virando os “editores” de conteúdo.

O tratamento de dados capaz de personalizar a informação que recebemos faz com que o Facebook e o Google acabem funcionando como guardiões (*gatekeepers*) de quais notícias e informações aparecem nos nossos *feeds* ou resultados de pesquisa. Agora, como esses algoritmos funcionam? Não se sabe aos detalhes. É claro que sabemos como eles funcionam tecnicamente, mas os algoritmos são propriedade privada e estão protegidos por estruturas legais de propriedade intelectual que bloqueiam algumas exigências de transparência. Em comparação, um jornal tradicional tem um editor que faz uma escolha de qual será a matéria principal. Embora nem tudo seja totalmente transparente na imprensa tradicional, esse editor e a política editorial desses órgãos podem ser questionados pela sociedade civil de uma forma que, pelo menos até agora, os algoritmos não podem ser.

Portanto, além das *fake news* serem mais lucrativas, não temos muito controle - nem se pode questionar muito - sobre como e por que uma ou outra notícia apareceu, em detrimento de outra.

---

<sup>41</sup> Tradução livre do autor: “A informação falsa difunde-se significativamente mais, mais depressa, mais profundamente e mais amplamente do que a verdade em todas as categorias de informação, e os efeitos foram mais intensos para notícias políticas falsas do que para falsas notícias sobre terrorismo, desastres naturais, ciência, lendas urbanas ou informações financeiras. Descobrimos que as notícias falsas eram mais inovadoras do que as notícias verdadeiras, o que sugere que as pessoas eram mais propensas a compartilhar informações novas.” in SOROUGH, Vosoughi. ROY, Deb. ARAL, Sinan. *The spread of true and false news online*, op. cit. nota 7.

Pois bem. Para tentar atenuar essa questão, recente artigo publicado por diversos pesquisadores da área propõe que as mídias sociais passem a ranquear as fontes de informação por um critério de qualidade:

As plataformas poderiam fornecer aos consumidores sinais da qualidade da origem da informação que poderiam ser incorporados aos rankings algorítmicos de conteúdo. Eles poderiam minimizar a personalização da informação política em relação a outros tipos de conteúdo (reduzindo a criação de “echo chambers”). As funções que enfatizam o conteúdo atual de tendências podem tentar excluir a atividade de bots das medições do que é tendência. Mais genericamente, as plataformas poderiam restringir a disseminação automatizada de conteúdo de notícias por bots e cyborgs (usuários que automaticamente compartilham notícias de um conjunto de fontes, com ou sem lê-los), embora no futuro previsível, os produtores de bots provavelmente serão capazes de projetar contramedidas eficazes.<sup>42</sup>

No entanto, permitir o ranqueamento de notícias pode se aproximar da figura de um curador da qualidade do discurso público, como se fosse possível situar algum ente (estatal ou privado) na posição de avaliar o que merece e o que não merece ser dito. De um lado, permitir o ranqueamento pelas plataformas pode levantar questões sérias sobre estarmos depositando poderes demais nessas empresas de tecnologia, que também possuem seus interesses econômicos e privados. Na outra face da moeda, caso soluções governamentais fossem tentadas, as mesmas questões surgem, pois se o Estado passasse a ranquear os órgãos da imprensa, possíveis “listas brancas” passariam a ser o problema. Daí para a censura e o controle dos discursos pelo governo faltaria muito pouco; e a liberdade de imprensa é crucial para um ambiente democrático. De um modo geral, vale a pena lembrar que as avaliações nas redes sociais costumam ser feitas pela própria sociedade civil, o que é certamente mais democrático do que listas de confiabilidade vindas de órgãos estatais.

---

<sup>42</sup> Tradução livre do autor do trecho “*The platforms could provide consumers with signals of source quality that could be incorporated into the algorithmic rankings of content. They could minimize the personalization of political information relative to other types of content (reducing the creation of “echo chambers”). Functions that emphasize currently trending content could seek to exclude bot activity from measures of what is trending. More generally, the platforms could curb the automated spread of news content by bots and cyborgs (users who automatically share news from a set of sources, with or without reading them), although for the foreseeable future, bot producers will likely be able to design effective countermeasures.*” in *The Science of Fake News*, op. cit. nota 5.

Portanto, embora seja inegável que as plataformas possam contribuir com soluções para o problema, o que elas já vêm fazendo, a sociedade civil também precisa, externamente, criar seus mecanismos de averiguação da verdade. Até porque o problema da disseminação de informações falsas não é culpa exclusiva de contas automatizadas.

O estudo do MIT sobre a disseminação de notícias falsas acima referido, confirma essa assertiva quando conclui que

o comportamento humano contribui mais para a disseminação diferencial de falsidades e verdades do que os robôs automatizados. Isso implica que as políticas de contenção de informações falsas também devem enfatizar intervenções comportamentais, como recompensar e incentivar a dissolução da disseminação de desinformação, em vez de se concentrar exclusivamente em cercear os bots. Entender como as notícias falsas se espalham é o primeiro passo para contê-las.<sup>43</sup>

Na toada de que as plataformas e a sociedade civil também precisam fazer o seu papel, é importante notar que uma série de medidas vêm sendo tomadas para contra-atacar o fenômeno. Cada vez mais surgem agências de checagem de fatos, ferramentas de conscientização dos usuários sobre *fake news* e de checagem de fatos por palavras-chaves e robôs automatizados, como a recente ferramenta brasileira PegaBot, ou as colunas e setores da mídia tradicional destinados apenas a desbancar informações falsas. Já do lado das plataformas, é crucial continuar a revisão dos termos de uso e das políticas de privacidade para dificultar a utilização de *bots* no impulsionamento de *fake news*, além de investir em inteligência artificial capaz de detectar conteúdo falso e ações orquestradas por *bots* e atores políticos maliciosos.

---

<sup>43</sup> Tradução livre do autor para o trecho “*human behavior contributes more to the differential spread of falsity and truth than automated robots do. This implies that misinformation-containment policies should also emphasize behavioral interventions, like labeling and incentives to dissuade the spread of misinformation, rather than focusing exclusively on curtailing bots. Understanding how false news spreads is the first step toward containing it. We hope our work in-spires more large-scale research into the causes and consequences of the spread of false news as well as its potential cures.*” in SOROUSH, Vosoughi. ROY, Deb. ARAL, Sinan. *The spread of true and false news online*, op. cit. nota 7.

## 1.2 O modelo russo de propaganda: guerra cibernética via desinformação: *troll armies* e *overload* informacional

Segundo as agências de inteligência norte-americanas, a Rússia interferiu ativamente nas eleições presidenciais de 2016 nos Estados Unidos por meio de campanhas de desinformação e propaganda realizadas em grande parte pelas redes sociais. Percebe-se, entretanto, que a interferência russa vem sendo notada em outros lugares, pois o modelo com características de guerra cibernética vem se repetindo em vários países ao longo dos últimos anos, especialmente à medida em que momentos eleitorais significativos se aproximam.<sup>44</sup> A influência em eleições se dá, resumidamente, pela distorção de fatos, pelo micro-direcionamento dos eleitores, para persuadi-los a votar de uma certa maneira, ou apenas para reprimir seu desejo de votar.

Um relatório da Universidade de Oxford publicado em julho de 2018 identificou evidências de campanhas organizadas de manipulação pela mídia social em 48 países.<sup>45</sup> As evidências levaram pesquisadores e agências de inteligência de alguns desses países a atribuírem à Rússia a tentativa de interferir na opinião pública, apoiando organizações que criam e disseminam desinformação, conteúdo falso e hiper-partidário, com o objetivo de minar a confiança pública e de desestabilizar os estados democráticos. Trata-se de uma ameaça ativa.

Este modelo de guerra cibernética pela via informacional pode ser analisado sob dois aspectos: em primeiro, pela propaganda computacional, comprovadamente aplicada<sup>46</sup> nos Estados Unidos,<sup>47</sup> Inglaterra,<sup>48</sup> França,

---

<sup>44</sup> Informação disponível em: <<https://www.newyorker.com/news/news-desk/what-muellers-indictment-reveals-about-russias-internet-research-agency>> (acessado em 04.02.19).

<sup>45</sup> BRADSHAW, Samantha. HOWARD, Phillip. N. *Challenging Truth and Trust: A Global Inventory of Organized Social Media Manipulation*. In: *The Computational Propaganda Project*. Oxford, United Kingdom: Oxford Internet Institute, University of Oxford, p. 1.

<sup>46</sup> Segundo o relatório "*Freedom on the Net 2017 - Manipulating Social Media to Undermine Democracy*" da Freedom House, em pelo menos 20 países, padrões característicos de atividade online sugeriram o uso coordenado de tais "bots" para influenciar o cenário político, op. cit. nota 37.

<sup>47</sup> Informação disponível em: <<https://www.nytimes.com/interactive/2018/02/16/us/politics/russia-propaganda-election-2016.html>> e <<https://www.nytimes.com/interactive/2018/02/16/us/politics/document-The-Special-Counsel-s-Indictment-of-the-Internet.html>> (acessados em 04.02.19).

México e Alemanha, e que também vem sendo notada no Brasil, já por alguns anos,<sup>49</sup> agindo de forma bastante sofisticada e eficaz, pelo agravamento da perseguição online abusiva, perpetrada por verdadeiros “exércitos cibernéticos” (também chamados de “*web brigades*” - Врегады - ou *troll armys*) que buscam ameaçar, hackear e ofender jornalistas e veículos de comunicação, para que pensem duas vezes antes de escrever conteúdo crítico, tornando o jornalismo político menos atraente.

Em suma, o modelo se utiliza da “arquitetura de controle das redes sociais” descrita nos tópicos anteriores para minar o ecossistema informacional, desviando a atenção direcionada a temas-chaves para debates políticos menos importantes por meio da manipulação de *trend topics* e criação estratégica de notícias falsas e cascatas cibernéticas, bem como pela perseguição da imprensa. Essas conclusões podem ser obtidas a partir de relatórios e artigos de instituições e pesquisadores do tema, os quais, por exemplo, conectam o Kremlin, através da “Agência de Pesquisa na Internet” (Internet Research Agency), a grupos ou organizações “sem fins lucrativos”, que operam nas redes sociais e na mídia, apontados como os responsáveis pela disseminação do modelo de guerra informacional russo.

Nessa toada, a recente notícia da Folha de São Paulo que denunciou<sup>50</sup> um suposto esquema de envio em massa de mensagens via Whatsapp programado para as últimas semanas das eleições está inserido no contexto, pois o modelo utilizado se parece em muito com as técnicas de desvio de atenção explicadas a seguir. Não obstante estas denúncias estarem ligadas à disseminação em massa de conteúdo nos últimos dias antes do segundo turno, já há indícios suficientes para confirmar que o modelo de propaganda vem tomando conta do ecossistema informacional brasileiro nos últimos anos, e é utilizado pelos mais diversos atores políticos - sem querermos, aqui, indicar um ou outro como responsável por qualquer “brasileirização” do modelo russo de propaganda.

---

<sup>48</sup> Informação disponível em: <<https://www.theguardian.com/uk-news/2018/feb/10/social-media-firms-must-tell-users-impacted-by-propaganda-mp-says>> (acessado em 04.02.19).

<sup>49</sup> Informação disponível em: <<http://blogs.oii.ox.ac.uk/politicalbots/wp-content/uploads/sites/89/2017/06/Comprop-Brazil-1.pdf>> (acessado em 04.02.19).

<sup>50</sup> Informação disponível em: <<https://www1.folha.uol.com.br/poder/2018/10/empresarios-bancam-campanha-contr-o-pt-pelo-whatsapp.shtml>> (acessado em 04.02.2019).

Conforme estudo do Instituto de Tecnologia e Sociedade do Rio de Janeiro (ITS-RIO) explicou,<sup>51</sup> o caso do ecossistema informacional brasileiro possui uma importante diferença de outras nações. No contexto brasileiro o uso do WhatsApp como canal de distribuição de *fake news* tem especial relevância porque o aplicativo - que já funciona em pelo menos 180 países –possui ao menos 120 milhões dos seus 1 bilhão de usuários no Brasil.<sup>52</sup> Além disso, o debate é especialmente importante para o nosso país porque pesquisas recentes indicam que, no Brasil, a grande maioria das pessoas recebe suas notícias por meio das redes sociais (cerca de 70%).<sup>53</sup> O uso do WhatsApp como canal de distribuição de notícias vem sendo amplamente noticiado<sup>54</sup> e pesquisado.

Não podemos esquecer, no entanto, que o WhatsApp, que também é controlado pelo Facebook, vem investindo em melhorias para evitar a desinformação em massa. Alguns dos exemplos de mudança na plataforma são: (i) a restrição a grupos para até 250 membros; (ii) a criação da advertência “encaminhada” quando mensagens são compartilhadas de uma janela para outra (sinalizando a possibilidade de ser uma corrente tentando propagar informação em cascata); (iii) restrições na quantidade de usuários em listas de transmissão e compartilhamento de mensagens.<sup>55</sup> Além disso, a empresa vem trabalhando em conjunto com o Judiciário - especialmente o Superior Tribunal Eleitoral. Mudanças como essas também foram implementadas em países como a Índia, algumas por razões similares, outras por particularidades sociais diferentes, como ondas de linchamentos públicos.<sup>56</sup>

---

<sup>51</sup> O recente estudo do ITS-RIO pode ser acessado em: <<https://feed.itsrio.org/poder-computacional-automa%C3%A7%C3%A3o-no-uso-do-whatsapp-nas-elei%C3%A7%C3%B5es-e969746d231f>> (acessado em 04.02.19).

<sup>52</sup> Informação disponível em: <<https://www.whatsapp.com/about/>> e em <<https://link.estadao.com.br/noticias/empresas,whatsapp-chega-a-120-milhoes-de-usuarios-no-brasil,70001817647>> (acessados em 04.02.19).

<sup>53</sup> Informação disponível em: <<http://observatoriodaimprensa.com.br/e-noticias/cerca-de-70-dos-brasileiros-se-informam-pelo-facebook/>> (acessado em 04.02.19).

<sup>54</sup> Sobre o tema ver: <<https://www.bbc.com/portuguese/brasil-45910249>> e <<https://www.bbc.com/portuguese/brasil-42172146>>.

<sup>55</sup> Sobre o tema ver o site do aplicativo de mensagens, especialmente o Blog que introduz as mudanças implementadas e as dicas de segurança, disponíveis respectivamente em: <<https://blog.whatsapp.com/>> e <<https://www.whatsapp.com/safety/>> (acessados em 04.02.19).

<sup>56</sup> Sobre o tema ver: <<https://www.bbc.com/news/world-asia-india-44897714>> (acessado em 04.02.19).

O assunto ganha ainda mais atualidade, pois em 18.10.2018, a apenas 10 dias do segundo turno, o Partido dos Trabalhadores (PT) ingressou com uma Ação de Investigação Judicial Eleitoral - AIJE (0601771-28.2018.6.00.0000) no Tribunal Superior Eleitoral (TSE), instando a corte a investigar denúncias de abuso de poder econômico e uso indevido de meios de comunicação digital por causa de suposta contratação de empresas de disparos de mensagens em massa - justamente no Whatsapp - que teria sido feita indiretamente por grupos econômicos apoiadores da candidatura de Bolsonaro. A conduta seria vedada pela lei eleitoral<sup>57</sup> e pela decisão do STF na ADIn 4650,<sup>58</sup> que proíbem a doação por empresas para campanhas eleitorais.

Nesse contexto, as pesquisas na área sugerem que este contemporâneo modelo de propaganda e de interferência na opinião pública internacional, com características de guerra cibernética, tem o potencial de ser altamente eficaz justamente porque se utiliza da maneira como o cérebro humano recebe, interpreta, armazena e compartilha informação, para direcionar o intérprete - na maioria das vezes o consumidor de informação online - aos objetivos pretendidos. O conteúdo produzido ao mesmo tempo que (i) entretém, (ii) confunde e (iii) sobrecarrega o público de notícias sobre um mesmo fato, (iv) angariando as primeiras atenções sobre o assunto devido ao fato de não promoverem checagem e por serem estrategicamente manipuladas, em seguida, (v) massacrando o público com repetições do conteúdo e outras (vi) técnicas retóricas e de design.

Em síntese, o sucesso do modelo de propaganda baseado em técnicas da Guerra Fria soviética envolve 4 características principais:

- 1) O grande volume de informação: múltiplos veículos e canais de comunicação
- 2) É rápido, contínuo e repetitivo
- 3) Não tem compromisso com a realidade objetiva
- 4) Falta de compromisso com a consistência (jornalística)

---

<sup>57</sup> Art. 22, inciso XIV, da Lei Complementar n. 64/90 e art. 323 do Código Eleitoral.

<sup>58</sup> Julgou procedente em parte o pedido formulado na Ação Direta de Inconstitucionalidade (ADI) 4650 para declarar a inconstitucionalidade dos dispositivos legais que autorizavam as contribuições de pessoas jurídicas às campanhas eleitorais.



Considerando que esses aspectos bem-sucedidos da propaganda russa têm fundamentos surpreendentes na literatura sobre psicologia e ciência comportamental, é necessário primeiro entender como funciona o modelo para traçar possíveis abordagens no mesmo campo, respondendo na medida, ou simplesmente competindo.

### 1.2.1 O grande volume de informação: múltiplos veículos e canais de comunicação

O modelo de que tratamos no item anterior produz um volume incrivelmente grande de informação, falsa, parcialmente falsa ou descontextualizada, que é transmitido ou distribuído através de um amplo número de canais: a propaganda foca em disseminar textos, vídeos, áudios e imagens, preferencialmente via Internet e redes sociais, como o Facebook e Twitter, e aplicativos de mensagens como o Whatsapp. No entanto, também pode se utilizar da televisão e radiodifusão.

O conteúdo das mensagens recorrentemente envolve notícias com pretensão verniz jornalístico, com manchetes apelativas, alegações chocantes, em formatos semelhantes aos veículos de imprensa autênticos, no entanto sem que haja a indicação de quem escreveu a reportagem ou fontes confiáveis. Notícias falsas frequentemente contêm imagens ou vídeos manipulados, ou até mesmo autênticos, mas descontextualizados.

Os produtores e disseminadores possuem uma rede substancial de atores políticos com centenas de grupos no Whatsapp - muitos dos quais as pessoas são adicionadas sem permissão -,<sup>59</sup> ou em outras plataformas como o Facebook e Twitter, contas falsas e bots.<sup>60</sup> No contexto específico de disputas político-eleitorais, os robôs podem ser manuseados para distorcer a dimensão de movimentos políticos, manipular e radicalizar debates na esfera pública, e criar falsas percepções sobre disputas e consensos nas redes sociais. Como explicado acima e em outro momento,<sup>61</sup> a prática denominada “*astroturfing*”

---

<sup>59</sup> No caso das eleições de 2018, foram noticiados diversos relatos de grupos de Whatsapp criados com o intuito de disseminar notícias, nos quais pessoas eram adicionadas sem seu consentimento. Sobre o assunto ver: <<https://www1.folha.uol.com.br/poder/2018/10/como-telefones-de-usuarios-do-facebook-foram-usados-por-campanhas-em-disparos-em-massa-no-whatsapp.shtml?loggedpaywall>>. Para saber como evitar que seu número de celular seja descoberto dentro do Facebook, ver <<https://www.bbc.com/portuguese/brasil-45922572>> (acessados em 04.02.19).

<sup>60</sup> Pra piorar o cenário, de acordo com estimativas do provedor de serviços na nuvem Imperva Incapsula, os robôs representaram 51,2% de todo o tráfego da web em 2016. Isso significa que, além da atividade humana na Internet e do uso saudável de robôs, há cada vez mais contas automatizadas nas mídias sociais para manipular discursos na esfera pública online. Disponível em: <<https://www.incapsula.com/blog/bot-traffic-report-2016.html>> (acessado em 04.02.19).

<sup>61</sup> Disponível em: <<https://www.jota.info/opiniao-e-analise/artigos/redes-sociais-conteudo-informacional-0608201>> (acessado em 04.02.19).

funciona da seguinte forma: múltiplas identidades online e grupos de pressão falsos com grande quantidade de bots são utilizados, por exemplo, para promover falsamente uma notícia com pretensão verniz jornalístico, introduzidos numa variedade de grupos políticos, criando a impressão de amplo apoio popular a uma ideia, política, indivíduo ou produto. Isso é atingido por meio da multiplicação do número de acessos e compartilhamentos potencializados por robôs programados com essas intenções e, paralelamente, por pessoas reais.

É interessante destacar que no Brasil, recente relatório<sup>62</sup> do InternetLab concluiu que no Twitter, entre aproximadamente 13 a 47% dos seguidores dos candidatos à Presidência são bots, demonstrando que a presença de bots no discurso político nas eleições de 2018 foi uma realidade.

Perceber essa prática com seus próprios olhos não é tão difícil: basta ver uma notícia importante de última hora, e perceber que, no Facebook, por exemplo, há comentários de contas bastante suspeitas<sup>63</sup> com uma enorme quantidade de likes e compartilhamentos. Entrando no perfil dessas contas, nota-se um perfil com pouquíssimo histórico de navegação, poucas informações disponíveis, fotos repetidas etc.

Voltando ao modus operandi do modelo de propaganda russo, especialmente na Internet, com o grande volume de mensagens e a quantidade de notícias falsas sobre um mesmo fato, de fato, as versões manipuladas saem na frente, pois em condições iguais de competição, a propaganda ofertada em maior volume e proveniente de uma maior quantidade de fontes se torna mais persuasiva. Primeiro, porque o grande volume pode consumir a atenção e desviar o foco de outros canais de informação disponíveis, abafando as mensagens concorrentes e desacreditando a mídia tradicional. Segundo, o alto volume pode sobrecarregar as mensagens concorrentes em uma inundação de discordância. Em terceiro lugar, com vários canais aumentam-se as chances de os públicos-alvo serem expostos à mensagem ou ideia disseminadas. Quarto, receber uma mensagem de vários lugares e de várias origens aumenta

---

<sup>62</sup> Disponível em: <<http://www.internetlab.org.br/wp-content/uploads/2018/07/Relat%C3%B3rio-Bots-ou-n%C3%A3o.pdf><https://www.jota.info/opiniao-e-analise/artigos/redes-sociais-conteudo-informacional-0608201>> (acessado em 04.02.19).

<sup>63</sup> O Facebook possui um setor específico dentro da empresa apenas para fazer a checagem de contas falsas, havendo uma opção específica para denunciar o comportamento na rede social.

a credibilidade percebida da mensagem, especialmente se uma fonte de divulgação for uma com a qual um membro da audiência se identifica, como amigos e familiares.

Em resumo, a variedade de fontes é importante porque:

- ↳ Múltiplas fontes são mais persuasivas do que uma única fonte, especialmente se essas fontes contiverem argumentos diferentes que apontem para a mesma conclusão.
- ↳ Receber a mesma mensagem ou mensagem similar de várias fontes é mais persuasivo.
- ↳ As pessoas supõem que as informações de várias fontes provavelmente se baseiam em perspectivas diferentes e, portanto, merecem uma consideração maior.

Nesse contexto, o número e quantidade de fontes também são importantes porque:

- ↳ O endosso de um grande número de usuários aumenta a confiança e a confiabilidade do consumidor na informação, muitas vezes com pouca atenção à credibilidade daqueles que fazem o endosso.
- ↳ Quando o interesse do consumidor é baixo, o poder de persuasão de uma mensagem pode depender mais do número de argumentos que o sustentam do que da qualidade desses argumentos.

Por fim, a opinião daqueles que estão mais próximos ao nosso círculo familiar e de amizades é crucial, porque:

- ↳ Comunicações de grupos aos quais o destinatário pertence têm maior probabilidade de serem consideradas credíveis. O mesmo se aplica quando a fonte é percebida como semelhante ao destinatário. Se um canal de propaganda é (ou pretende ser) de um grupo com o qual o receptor se identifica, é mais provável que seja persuasivo.
  
- ↳ A credibilidade pode ser social, ou seja, é mais provável que as pessoas percebam uma fonte como credível se outras pessoas perceberem a fonte como confiável. Esse efeito é ainda mais forte quando não há informações suficientes disponíveis para avaliar a confiabilidade da fonte.
  
- ↳ Quando o volume de informações é baixo, os destinatários tendem a favorecer os especialistas, mas quando o volume de informações é alto, os destinatários tendem a favorecer as informações de outros usuários.

### 1.2.2 O modelo é rápido, contínuo e repetitivo

O modelo de propaganda russo tem a agilidade de ser o primeiro a produzir a notícia, o que permite aos propagandistas a oportunidade de criar a primeira impressão. Junto a isso, a combinação de mensagens em grande volume, multicanais e contínuas tornam os temas escolhidos mais familiares ao seu público, o que lhes dá um impulso em termos de credibilidade, experiência e confiabilidade.

Devido à sua falta de compromisso com a realidade objetiva - que será discutida mais adiante -, eles não precisam esperar para verificar os fatos ou alegações, como costumeiramente fazem os veículos tradicionais de informação; eles apenas disseminam uma interpretação de eventos emergentes que parecem favorecer melhor seus temas e objetivos. Isso os permite ser notavelmente responsivos e ágeis, muitas vezes transmitindo as primeiras “notícias” dos eventos (e, com frequência similar, as primeiras notícias de não eventos, ou coisas que na verdade não aconteceram). Os canais repetem e reciclam a desinformação entre si.

O que piora essa situação é que por vezes esse *overload* de informação é captada e retransmitida por meios de comunicação legítimos. Com frequência, a mídia social convencional repete os temas, mensagens ou falsidades introduzidas por um dos muitos canais de *fake news*. Nesse contexto, é mais provável que um indivíduo aceite as primeiras informações recebidas sobre um assunto e, em seguida, se agarre à essas informações quando confrontado com mensagens conflitantes. Além disso, a repetição leva à familiaridade e a familiaridade leva à aceitação. Então:

- ↳ A exposição repetida a uma declaração tende a aumentar sua aceitação como verdadeira.
- ↳ O “efeito de verdade ilusório” está bem documentado, segundo o qual as pessoas classificam as declarações como mais

verdadeiras, válidas e verossímeis quando já encontraram essas declarações antes do que quando são novas declarações.

↳ Quando as pessoas estão menos interessadas em um tópico, elas são mais propensas a aceitar a familiaridade trazida pela repetição como um indicador de que a informação (repetida ao ponto de familiaridade) está correta.

↳ Ao processar informações, os consumidores podem economizar tempo e energia usando uma heurística de frequência, ou seja, favorecendo as informações que ouviram com mais frequência.

↳ Mesmo com histórias absurdas e lendas urbanas, aqueles que as ouviram várias vezes são mais propensos a acreditar que são verdadeiras.

↳ Se um indivíduo já estiver familiarizado com um argumento ou reivindicação (já o viu antes, por exemplo), ele o processará com menos cuidado, muitas vezes deixando de discriminar argumentos fracos de argumentos fortes.

### 1.2.3 O modelo não tem compromisso com a realidade objetiva

Muito embora o conceito de “verdade” seja algo relativo, existe um padrão objetivo de realidade factual ou ao menos padrões éticos de jornalismo e de redação, costumeiramente seguidos pela mídia convencional.

Por outro lado, além de fabricar informações, o modelo do qual estamos falando também fabrica fontes e material que supostamente serviria como “prova”. Quem nunca recebeu uma *fake news* com supostas imagens e vídeos, refutou e escutou um questionamento do tipo: “*Então esse vídeo é armado*”? Justamente.

Em tese, a propaganda tem pouco ou nenhum compromisso com a verdade, mas isso não quer dizer que todo o conteúdo dito seja falso. Pelo contrário, muitas vezes o conteúdo tem uma fração significativa de verdade. Às vezes, no entanto, eventos relatados são totalmente fabricados. Algumas das imagens que acompanham as *fake news* são posteriormente expostas como falsas devido à má edição de fotos, como no recente caso das urnas que supostamente estavam com defeitos.<sup>64</sup>

*Fake news* são aceitas em parte porque as pessoas simplesmente não as reconhecem como falsas ou porque várias pistas levam-nas a atribuir-lhes maior credibilidade do que deveriam. Portanto, quanto mais sofisticada for a veiculação da notícia falsa, mais chance ela tem de ser acreditada e propagada em cascatas cibernéticas.<sup>65</sup>

Outro problema deve ser notado: as informações falsas - em especial as que apelam para as emoções - são mais lucrativas do que links com notícias verdadeiras. Ora, é óbvio que as mídias sociais precisam do dinheiro de publicidade e, como explicamos acima, nos modelos de negócio praticados por essas plataformas, esse dinheiro é gerado por cliques, de forma que as *fake*

---

<sup>64</sup> Disponível em: <<http://www.tre-mg.jus.br/imprensa/noticias-tre-mg/2018/Outubro/justica-eleitoral-esclarece-boato-sobre-processamento-dos-votos-na-urna-antes-da-tecla-confirma>> (acessado em 04.02.19).

<sup>65</sup> Sobre as cascatas cibernéticas ver: SUNSTEIN, Cass R. #Republic: *Divided Democracy in the Age of Social Media*, op. cit. nota 10. Capítulo. 4.



news conseguem mais cliques e ganham mais dinheiro do que as notícias verdadeiras.<sup>66</sup>

Além disso, as “matérias” e sites são visualmente parecidos com os de noticiários convencionais, e as pessoas que aparecem neles são representados como jornalistas, especialistas, ou famosos ativistas em redes sociais, tornando os membros da audiência muito mais propensos a atribuir credibilidade à desinformação. Percebe-se também que a proporção das falsidades aumenta ainda mais quando a desinformação é consistente com narrativas ou preconceitos mantidos pelo público - preconceitos não raros no Brasil -, e direcionada a públicos-alvo específicos, mais propensos a aceitar as ideias as quais se busca introduzir.

É sob esse enfoque que o recente caso Cambridge Analytica chocou o mundo, pois, por meio do tratamento dos dados pessoais de cada usuário coletados no Facebook, a empresa conseguiu desenvolver perfis “psicográficos” para cada um e permitir um direcionamento de discurso ainda mais subjetivo. Individualizando-se ao máximo os consumidores, as plataformas são capazes de definir nossas preferências de consumo e “empurrar” produtos e serviços,<sup>67</sup> ou, conteúdo político estrategicamente selecionado.

Por que essa desinformação pode ser eficaz? Primeiro, as pessoas costumam ser cognitivamente preguiçosas. Devido à sobrecarga de informação (especialmente na Internet), elas usam uma série de heurísticas e atalhos diferentes para determinar se as novas informações são confiáveis. Segundo, as pessoas geralmente são ruins em discriminar informações verdadeiras de informações falsas - ou lembrar-se de que já fizeram isso anteriormente, no primeiro contato com a mentira. Temas familiares ou mensagens podem ser

---

<sup>66</sup> Como o estudo do MIT indicou, em tradução livre do autor: “A informação falsa difunde-se significativamente mais, mais depressa, mais profundamente e mais amplamente do que a verdade em todas as categorias de informação, e os efeitos foram mais intensos para notícias políticas falsas do que para falsas notícias sobre terrorismo, desastres naturais, ciência, lendas urbanas ou informações financeiras. Descobrimos que as notícias falsas eram mais inovadoras do que as notícias verdadeiras, o que sugere que as pessoas eram mais propensas a compartilhar informações novas.” in SOROUGH, Vosoughi. ROY, Deb. ARAL, Sinan. *The spread of true and false news online*, op. cit. nota 7.

<sup>67</sup> Sobre o tema ver: THALER, R. H., SUNSTEIN, C. R. *Nudge: improving decisions about health, wealth, and happiness*. Rev. and expanded ed. New York: Penguin Books. (2009). Ver também: SUNSTEIN, C. R. *Human Agency and Behavioral Economics*. Cham, Switzerland: Palgrave Macmillan, 2017.

atraentes, mesmo se esses temas e mensagens forem falsos. Informações que se conectam com identidades de grupo ou narrativas familiares - ou que despertam emoção - podem ser particularmente persuasivas.

Finalmente, a credibilidade da fonte é frequentemente avaliada com base em “pistas periféricas”, que podem ou não estar de acordo com a realidade da situação. Um vídeo, por exemplo, que parece uma transmissão de notícia, mesmo que seja na verdade uma transmissão de propaganda, pode conferir o mesmo grau de credibilidade que um noticiário real. Descobertas do campo da psicologia mostram como as pistas periféricas podem aumentar a credibilidade da propaganda. Em suma, os efeitos dessa abordagem são:

- ↳ Em um fenômeno conhecido como “*sleeper effect*”, fontes de baixa credibilidade manifestam maior impacto persuasivo com o passar do tempo. Enquanto as pessoas fazem avaliações iniciais da credibilidade de uma fonte, ao lembrarem do tema em outra ocasião, a informação é frequentemente dissociada de sua fonte. Assim, a informação de uma fonte questionável pode ser lembrada como verdadeira, embora a fonte seja esquecida.
- ↳ É mais provável que alguém aceite informações quando essas são consistentes com outras mensagens que a pessoa acredita serem verdadeiras (assuntos conexos).
- ↳ As pessoas sofrem de “*confirmation bias*” (viés de confirmação): elas veem notícias e opiniões que confirmam as suas crenças já existentes como mais confiáveis do que outras notícias e opiniões, independentemente da qualidade dos argumentos.
- ↳ Pessoas cujo grupo de pares (familiares, amigos e conhecidos) é afetado por um evento são muito mais propensas a aceitar teorias de conspiração sobre esse evento.

- ↳ Histórias ou relatos que criam excitação emocional no receptor (por exemplo, repugnância, medo, felicidade) são muito mais passíveis de serem transmitidos, sejam verdadeiros ou não.
- ↳ Mensagens furiosas são mais persuasivas para o público irritado.
- ↳ A presença de suposta evidência pode anular os efeitos da credibilidade da fonte na veracidade percebida das declarações.
- ↳ Em simulações de Tribunais, testemunhas que fornecem mais detalhes - até detalhes triviais - são consideradas mais confiáveis.
- ↳ Dicas periféricas, como a aparência de especialização, são as principais dimensões de credibilidade e essas qualidades podem ser avaliadas com base em pistas visuais, como formato, aparência ou simples declarações de especialização, levando as pessoas a aceitar - com pouca reflexão - que as informações vêm de uma fonte confiável.
- ↳ Sites de noticiários online são percebidos como mais confiáveis do que outros conteúdos online, como blogs de jornalistas, independentemente da veracidade do conteúdo.

#### 1.2.4 Falta de compromisso com a consistência jornalística

A característica distintiva final do modelo de propaganda russo é que ele não está comprometido com a consistência: os atores políticos simplesmente não têm problema em possuir duas ou mais versões (ou opiniões) sobre um mesmo fato. Mais uma vez, isso contradiz a sabedoria convencional sobre influência e persuasão. Se as fontes não forem consistentes, como elas podem ser confiáveis? Se elas não são credíveis, como podem ser influentes?

Primeiro, diferentes meios de propaganda não transmitem necessariamente os mesmos temas ou mensagens. Segundo, diferentes canais não transmitem necessariamente a mesma versão de eventos contestados. Terceiro, diferentes canais ou representantes não mostram medo de “mudar de tom”. Se uma falsidade ou deturpação é exposta ou não é bem recebida, os propagadores simplesmente a descartarão e passarão para uma nova, embora não necessariamente mais plausível.

Pesquisas sugerem que a inconsistência pode ter efeitos deletérios na persuasão - por exemplo, quando os destinatários fazem um esforço para examinar mensagens inconsistentes da mesma fonte. Mas com a rapidez e troca de informações, com qual regularidade as pessoas, de fato, checam por inconsistências? Além do mais, a literatura em psicologia experimental também mostra que as audiências podem ignorar as contradições sob certas circunstâncias:

- ↳ As contradições podem levar os destinatários a buscarem a entender porque ocorreu uma mudança na opinião. Quando um argumento aparentemente forte para uma mudança é fornecido ou assumido (por exemplo, mais pensamento foi dado ou mais informação foi obtida), a nova mensagem pode ter um maior impacto persuasivo.

↳ Quando uma fonte parece ter considerado diferentes perspectivas, a confiança de atitude do consumidor é maior. Uma fonte que muda sua opinião ou mensagem pode ser percebida como tendo dado maior consideração ao tópico, influenciando, assim, a confiança do destinatário na mais nova mensagem.

Além do mais, potenciais perdas na credibilidade devido à inconsistência são potencialmente compensadas por sinergias com outras características da propaganda contemporânea. Como observado anteriormente na discussão da multiplicidade de canais, a apresentação de múltiplos argumentos por múltiplas fontes é mais persuasiva do que a apresentação de múltiplos argumentos por uma fonte ou a apresentação de um argumento por múltiplas fontes. Em paralelo, essas perdas também podem ser compensadas por indicações periféricas que reforçam percepções de credibilidade, confiabilidade ou legitimidade, como a aparência da informação e os vídeos e imagens que as acompanham.

Mesmo se um canal ou propagandista individual mudar a versão de eventos de um dia para o outro, os espectadores provavelmente avaliarão a credibilidade da nova versão sem dar muito peso ao conteúdo da antiga, desde que haja novos sinais periféricos sugerindo que a fonte buscou mais e novas informações.

O problema ainda se acentua: apesar de posterior *fact checking*, as notícias falsas persistem. Não só as notícias verificadoras não chegam às mesmas pessoas que visualizam as *fake news* - uma vez que cada uma delas se encontra em sua bolha de filtragem algorítmica, que tende a rejeitar conteúdo (ou páginas e sites) das quais o usuário divirja -, mas a reiteração das notícias “verdadeiras” pode conceder ainda mais credibilidade às falsas.

A conclusão, então, é alarmante: as *fake news* e os falsos rumores, que são um dos mecanismos utilizados pelo modelo russo de propaganda, estão atingindo mais pessoas, penetrando mais fundo na rede social e se espalhando com mais rapidez do que as histórias precisas. Em contraste, a “checagem de fatos” pode acabar parecendo uma minoria e, portanto, uma opinião menos crível, resultando no mais devastador problema dos tempos de

pós-verdade: a desvalorização e deslegitimação de vozes especialistas, instituições autorizadas e o conceito de dados objetivos. Tudo isso acaba minando a capacidade da sociedade de se envolver em um discurso racional baseado em fatos compartilhados, deturpando o que seria uma visão majoritária, e o funcionamento da esfera pública.

## 2. PENSANDO UM MODELO DE ESFERA PÚBLICA DIGITAL

Tendo em vista que as Sociedades Algorítmicas possuem como característica a coleta de uma vasta quantidade de dados sobre os indivíduos, criando novas formas de vigilância e controle, tanto por governos quanto por empresas privadas, mostra-se fundamental aprofundar o debate sobre a ampliação da esfera pública para o mundo digital; isto é, a conceptualização da esfera pública clássica não é mais suficiente para compreender o *locus* - agora com sentido ainda mais intangível, e certamente muito mais complexo, dinâmico e veloz - no qual as relações sociais se travam.

Cass Sunstein, Professor da Faculdade de Direito da Universidade de Harvard, já havia alertado que essa arquitetura de controle (*architecture of control*) tem suas desvantagens, levantando questionamentos fundamentais sobre liberdade, democracia e autogoverno. Em sua obra “*#Republic: Divided Democracy in the Age of Social Media*” (2017), Sunstein faz um apelo por serendipidade (*serendipity*), que, em outras palavras, significa um apelo pelo acaso.

Em contraste com a ideia de uma sociedade na qual recebemos conteúdos de forma personalizada, como na atual Sociedade Algorítmica, Sunstein argumenta pela importância dos encontros casuais (*chance encounters*) e das experiências compartilhadas (*shared experiences*) como requisitos essenciais de sociedades democráticas, por causa do contato com a pluralidade e pela cultura democrática que acabam promovendo. Sem dúvida, essas características estão dentro da noção clássica de “fóruns públicos”, que é, por sua vez, um conceito largamente utilizado pelos estudos sobre democracia e esfera pública, e que igualmente será abordado aqui.

Quais são as pré-condições sociais para que um sistema de deliberação democrática funcione bem ou para a própria liberdade individual? Pode o acaso ser importante, mesmo que as pessoas não o queiram? Um universo de comunicações perfeitamente controlado - um feed personalizado - poderia ser seu próprio tipo de distopia? Como as mídias sociais, a explosão das opções de comunicação, o

aprendizado de máquina e a inteligência artificial podem alterar a capacidade dos cidadãos de se governarem?<sup>68</sup>

É necessário deixar claro que, ao debater as consequências negativas da personalização da Internet, os pesquisadores não estão idilicamente buscando um passado utópico, como se houvesse um ar nostálgico em relação à esfera pública pré algoritmos de relevância e busca. Na verdade, nota-se que, especialmente com respeito ao ecossistema informacional e às telecomunicações, o passado não era idílico e nem assim enxergado. Comparado a qualquer outro momento da história humana, estamos no meio de um período de ganhos extraordinários em relação à comunicação e igualmente do ponto de vista da própria democracia, pois, nunca antes - tecnologicamente falando - fomos tão capazes de promover princípios basilares como a pluralidade de discursos, o acesso à informação, a transparência etc. É bom que fique claro que, pelo menos em relação a isso, as coisas estão melhorando, e não piorando.

No entanto, essa Sociedade Algorítmica que coleta, analisa, usa e controla dados para promover políticas, sejam políticas públicas pelo Estado, sejam ofertas de produtos e serviços pelo mercado, acaba criando novos tipos de relações sociais e uma economia de poder diferente. Sob um ponto de vista, o desenvolvimento tecnológico trouxe às pessoas a esperança de que as novas tecnologias superem hierarquias políticas e sociais, distribuindo o poder político, organizacional e de comunicação, porém, de outro, é possível afirmar que, de fato, a Internet possibilitou uma maior distribuição do poder comunicacional, mas ele não necessariamente se democratizou, pois, há uma distinção fundamental entre poder distribuído e poder democrático. Jack Balkin, Professor da Faculdade de Direito da Universidade de Yale, ensina que, se por um lado o poder é democrático quando muitas pessoas participam da tomada

---

<sup>68</sup>SUNSTEIN, Cass R. *#Republic: Divided Democracy in the Age of Social Media*, op.cit. Nota 10. p. 5. Tradução livre do autor para o trecho original: “*What are the social preconditions for a well-functioning system of democratic deliberation or individual liberty itself? Might serendipity be important, even if people do not want it? Might a perfectly controlled communications universe—a personalized feed—be its own kind of dystopia? How might social media, the explosion of communications options, machine learning, and artificial intelligence alter the capacity of citizens to govern themselves?*”.



de decisão, por outro ele é distribuído quando opera em muitos lugares diferentes e afeta a uma multiplicidade de pessoas e situações distintas.<sup>69</sup>

Atualmente, a coleta de dados e sua manipulação estão amplamente distribuídas, mas isso não significa dizer que esse poder seja democraticamente controlado, porque, apesar dos dados serem coletados em vários lugares, apenas um número relativamente pequeno de organizações (e pessoas) detém os recursos e habilidades técnicas para coletá-los, tratá-los e usá-los. Entender essa assimetria é crucial para entender as mudanças ocorridas na esfera pública.

Para as finalidades deste capítulo, em primeiro lugar, aborda-se brevemente o conceito de esfera pública clássica, ou seja, como é concebida a noção de esfera pública por pensadores-chaves, ressaltando como o desenvolvimento tecnológico, em especial as novas formas de comunicação, alteram a concepção clássica do conceito de esfera pública. Neste momento, são tecidos breves comentários sobre a importância da liberdade de expressão e da deliberação na esfera pública e o que elas visam promover: a democracia e a qualidade das decisões políticas, respectivamente.

A partir daí, compara-se o contexto normativo habermasiano de esfera pública burguesa, que moldou a conceptualização clássica de esfera pública, à noção ainda em construção de esfera pública digital. Neste ponto, ressalta-se que na esfera clássica, o acesso aos veículos de comunicação, em termos gerais, era privilégio de uma elite burguesa intelectual, enquanto o que acontece na esfera pública digital é a ascensão das subjetividades de literalmente qualquer indivíduo conectado. Ademocratização do debate, portanto, afeta os parâmetros fundamentais da esfera pública, pois blogueiros e usuários de redes sociais não estão sujeitos a qualquer forma de controle,

---

<sup>69</sup> “Through its technologies of information collection, use, analysis, and control, the Algorithmic Society creates a new economy of power. The early Internet gave people hope that new technology would level political and social hierarchies by spreading political, organizational, and communicative powers widely among the world’s population. People hoped that the Internet and its related technologies would be a democratizing force. To some extent, that has been the case. But it is not quite how things have turned out. We should make a key distinction between distributed and democratic power. A form of power is democratic if many people participate in it and participate in decision making about how to employ it. A form of power is distributed if it operates in many different places and affects many different people and situations. In some ways the Internet and its associated digital technologies have made power more democratic. But in other ways the Internet has made it possible for power to be widely distributed but not democratic.” in BALKIN, Jack M. *Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation*, op. cit. nota 8. p. 7.

como padrões ético-jornalísticos e editoriais (*gatekeeping*) mínimos, fenômeno que também pode ser notado pela ascensão do amadorismo e da simplificação dos debates. O objetivo aqui é, resgatando o que foi dito na primeira parte do trabalho, demonstrar que a diminuição nas experiências comuns e um sistema de filtragem individualizada do conteúdo informacional que recebemos, bem como a ascensão do amadorismo e da simplificação do debate, alteraram o conceito clássico de esfera pública, podendo comprometer o próprio ideal de democracia no qual o conceito foi concebido, a ideia de deliberação, ou até a noção de uso público da razão nos espaços de deliberação.<sup>70</sup>

---

<sup>70</sup> Como diz John Rawls, em seu último escrito intitulado “The Idea of Public Reason Revisited”, de 1997: “A ideia de razão pública explicita no nível mais profundo os valores morais e políticos que devem determinar a relação de um governo democrático constitucional com os seus cidadãos e a relação destes entre si. (...) A ideia de razão pública tem uma estrutura definida e, se um ou mais dos seus aspectos é ignorado, ela pode parecer implausível, como parece quando aplicada à cultura de fundo. Ela tem cinco aspectos diferentes: (1) as questões políticas fundamentais às quais se aplica; (2) as pessoas a quem se aplica (funcionários do governo e candidatos a cargo público); (3) seu conteúdo como dado por uma família de concepções políticas razoáveis de justiça; (4) a aplicação dessas concepções em discussões de normas coercitivas a serem decretadas na forma de lei legítima para um povo democrático; (5) a verificação pelos cidadãos de que os princípios derivados das suas concepções de justiça satisfazem o critério de reciprocidade” (RAWLS, John. *O Direito dos Povos*. São Paulo: Martins Fontes, 2001, p. 174-175).

## 2.1. A esfera pública clássica e a esfera pública digital: semelhanças e diferenças

Segundo o *Oxford Bibliographies*, “a ‘esfera pública’ é geralmente concebida como o espaço social no qual opiniões diferentes são expressas, problemas de interesse geral são discutidos e soluções coletivas são desenvolvidas de forma comunicativa.”<sup>71</sup>

Certamente, a esfera pública foi um dos mais populares e mais debatidos conceitos da filosofia política do século XX, sendo possível dizer que é de certa forma estudada por todas as ciências sociais, uma vez que representa a arena central da comunicação social, aquela onde são transmitidas informações, sendo igualmente o espaço no qual os indivíduos deliberam entre si.

O termo “esfera pública” é uma tradução da palavra alemã “*Öffentlichkeit*”, expressão que se traduz em dois termos relacionados, porém distintos: “o público”, ou o coletivo de falantes e ouvintes presentes na esfera pública, e “publicidade”, melhor explicado como o estado de ser publicamente visível e sujeito a escrutínio pelo público.

Embora possa ser conceituada, a natureza controversa do conceito de esfera pública já é notada desde o início do século XX, não havendo equívoco por parte de quem afirma que as pesquisas modernas na esfera pública foram desencadeadas pela e a partir da obra “*Strukturwandel der Öffentlichkeit*” de Jürgen Habermas, publicada em alemão em 1962 e traduzida para o inglês apenas em 1989, posterior à primeira tradução em língua portuguesa de 1984.<sup>72</sup> Neste trabalho, escrito como tese de pós-doutorado exigida para

---

<sup>71</sup> No caso, o verbete “Public Sphere”, do *Oxford Bibliographies*, foi escrito pelos Professores Hartmut Wessler e Rainer Freudenthaler da Universidade de Mannheim. Tradução livre do autor para o trecho: “*The ‘public sphere’ is generally conceived as the social space in which different opinions are expressed, problems of general concern are discussed, and collective solutions are developed communicatively*”. Disponível em: <<http://www.oxfordbibliographies.com/>> (acessado em 04.02.19).

<sup>72</sup> HABERMAS, Jürgen. *Strukturwandel der Öffentlichkeit*. Neuwied/Berlim: Luchterhand, 1962 (tradução inglesa: *The Structural Transformation of the Public Sphere*. Cambridge, MA: MIT Press, 1989; tradução portuguesa: *Mudança Estrutural da Esfera Pública*. Rio de Janeiro: Tempo Brasileiro, 1984). Uma reedição da obra foi publicada em 1990, com novo e longo prefácio de Habermas, sobre a qual se baseia a segunda tradução em língua portuguesa (São Paulo: Editora UNESP, 2014, 565 p.).

obtenção do cargo de professor na Alemanha e dedicado às investigações sobre uma categoria da sociedade burguesa, como nota no subtítulo da obra, Habermas apresentou seu agora famoso argumento sobre a ascensão e decadência da esfera pública, à medida que ocorre a substituição de uma discussão pública mediante razões, em espaços públicos de sociabilidade, cujo resultado seria a formação de uma opinião pública ou um interesse comum, pela prática da agregação, negociação e compromisso de interesses privados no seio das democracias de massa cuja dinâmica da esfera pública se rende às sofisticadas técnicas dos novos meios de comunicação, atribuindo uma aura de prestígio às autoridades públicas, transformando a política em espetáculo dirigido e excluindo a maioria da população da discussão pública e dos processos de tomada de decisões.

No artigo “*The public sphere as an essentially contested concept: A co-citation analysis of the last 20 years of public sphere research*”,<sup>73</sup> texto cujo objetivo era o de revisar pesquisas recentes na esfera pública por meio de uma análise quantitativa de “co-citação”, Adrian Rauchfleisch, Professor da Universidade de Zurique, explica que o conceito de esfera pública não é fixo nem estável, como é típico para um conceito essencialmente contestado e construído por diferentes pensadores observando acontecimentos e momentos históricos distintos. A análise quantitativa de citações feita pelo pesquisador ajuda a obter uma visão mais ampla do desenvolvimento do conceito em diferentes disciplinas, para traçarmos a linha mestra que nos interessa.

Tal abordagem se mostra de extrema valia também para os fins deste capítulo, porque, em vez de encontrar a interpretação mais proeminente ou mesmo “correta” do conceito, o objetivo principal deste tópico é explorar, em linhas gerais, a evolução e a diferenciação do conceito antes e depois das transformações tecnológicas que afetaram drasticamente o ecossistema informacional nas duas últimas décadas, uma vez que na forma como a informação é disseminada, a mídia e a comunicação são consideradas epicentros do debate sobre a esfera pública.

---

<sup>73</sup> RAUCHFLEISCH, Adrian. “The public sphere as an essentially contested concept: A co-citation analysis of the last 20 years of public sphere research”. *Communication and the Public*, 2/1 (2017): 3-18.

No artigo de Rauchfleisch há uma interessante divisão das condições que caracterizam um conceito essencialmente contestado, exatamente como o caso do conceito de esfera pública, valendo a pena mencioná-las para entender como e por que, a partir de cada uma dessas condições e características, o conceito de esfera pública foi se desenvolvendo de modo orgânico e colaborativo.

A primeira condição é que um conceito essencialmente contestado seja (i) avaliativo (*appraisiveness*). Por ser avaliável, um conceito essencialmente contestado possui “valência normativa vinculada” a ele. Por exemplo, ao introduzir a esfera pública burguesa como a versão ideal para entender sua concepção, Habermas acaba anexando ao conceito uma alta carga normativa, uma vez que sua ideia mais importante é a do princípio crítico da esfera pública, isto é, a ideia de que as opiniões pessoais dos indivíduos privados podem desenvolver-se num processo de debate racional crítico, aberto a todos e livre de dominação, em que prevalece a força do melhor argumento, tornando o princípio do uso público da razão o conceito nuclear de uma teoria da democracia ou da política deliberativa.<sup>74</sup> Indo além, por exemplo, hoje em dia não há dúvidas de que a esfera pública é vista, em uma visão bastante pragmática, como uma pré-condição normativa para a democracia, pelo menos no contexto ocidental. Nota-se, então, a alta carga normativa do conceito.

A segunda e terceira condições são (ii) a complexidade interna (*internal complexity*) do conceito e (iii) sua descrição diversificada (*diverse describability*). Estas seriam condições inter-relacionadas, pois a complexidade interna de um conceito torna provável que pensadores diferentes possam ver e ordenar seu significado de maneiras distintas. Assim, por ser um conceito complexo com termos gerais genéricos, permite a descrição diversificada, uma

---

<sup>74</sup> Como, por exemplo, na seguinte passagem: “Se outrora os interesses individuais, neutralizados, visto que eram privatizados sob o denominador comum do interesse de classe, puderam atribuir uma certa racionalidade e também uma efetividade à discussão pública, hoje aparece em seu lugar a demonstração de interesses concorrentes. O consenso produzido na discussão pública mediante razões cede ao compromisso não público conquistado ou simplesmente imposto. As leis produzidas por essa via, mesmo quando o momento da generalidade é preservado em muitos casos, não podem mais reivindicar o momento de ‘verdade’, pois também foi despedaçada a esfera pública parlamentar, o lugar onde ela teria que se comprovar” (HABERMAS, Jürgen. *Mudança Estrutural da Esfera Pública*. São Paulo: Editora UNESP, 2014, p. 392).

vez que cada autor, a partir da sua cultura e de seu enfoque científico, oferece descrições únicas e distintas.

A quarta condição seria sua (iv) abertura (*openness*). Aqui, deve-se levar em conta que o caráter espacial e temporal da produção de conceitos sempre foi enxergado como um aspecto central de sua abertura, em razão de que, com a mudança de circunstâncias fáticas, um conceito também acaba precisando ser modificado e readequado, deixando um espaço aberto para sua evolução. Nessa perspectiva, considerando as incertezas envolvidas quanto aos desenvolvimentos futuros, o conceito de esfera pública certamente cumpre essa condição, porque suas modificações não puderam (e não podem) ser prescritas ou previstas com antecedência.

Por exemplo, Habermas, analisou a esfera pública na década de 1960, com foco na mídia de massa existente. Naquela época, a compreensão da esfera pública era muito engessada aos meios existentes para se transmitir informação. Por outro lado, acadêmicos contemporâneos já sabem que os limites da esfera pública não são fixos, mas se expandem e contraem com o tempo, de maneira que, para o que importa a este estudo, deve-se considerar que a televisão e a Internet tiveram um enorme impacto no conceito, comprovando sua abertura. Não restam dúvidas, portanto, que o tempo no qual se está avaliando o conceito de esfera pública é fundamental para sua concepção, motivo pelo qual qualquer análise da esfera pública digital das Sociedades Algorítmicas precisa levar em conta a arquitetura de controle das plataformas.

A quinta condição, talvez uma das mais interessantes para estudos em que o método comparativo com outros sistemas normativos é utilizado - como neste trabalho -, é a do (v) reconhecimento recíproco (*reciprocal recognition*) do conceito. Rauchfleisch lembra que o conceito habermasiano de esfera pública foi debatido criticamente por influentes pensadores da teoria feminista, ao exemplo de Seyla Benhabib e Nancy Fraser, bem como por estudiosos da tradição liberal americana, como Michael Schudson, colaboradores, ao lado de outros proeminentes teóricos, daquela obra que, após "*Strukturwandel der*

*Öffentlichkeit e Faktizität und Geltung*”(1992)<sup>75</sup> de Habermas, é a mais citada na literatura sobre esfera pública: o volume coletivo “*Habermas and the Public Sphere*” (1992), editado por Craig Calhoun.<sup>76</sup> Fraser e Benhabib questionaram diretamente a concepção de Habermas e, ao mesmo tempo, promoveram sua compreensão.<sup>77</sup> Ao analisar o conceito posto e criticá-lo, há, no mínimo, o reconhecimento de sua existência. O próprio Habermas, já no texto ao final do volume,<sup>78</sup> oferece ampla revisão de seu argumento original, apesar de manter sua perspectiva crítica e normativa da esfera pública, reconhecendo não poucas críticas e incorporando-as em seus escritos posteriores. Ambos os lados se reconhecem e a disputa em torno do conceito permanece em aberto.

A sexta condição consiste na utilização de (vi) exemplares (*exemplars*). Um conceito essencialmente contestado ancora-se em um exemplar original cuja autoridade é aceita por todos, ou quase todos, os estudiosos. No caso da esfera pública, esse exemplar original pode ser identificado, pois a investigação histórica e a preferência normativa sobre a esfera pública burguesa na França e na Inglaterra dos séculos XVIII e XIX, tal como aparece na obra *Strukturwandel der Öffentlichkeit* de Habermas, tornou-se o principal exemplar do conceito de esfera pública. Não que não haja contestação, a começar pela teoria feminista supramencionada, mas, como nota Rauchfleisch,<sup>79</sup> essa compreensão é mencionada como referência em estudos recentes, inclusive no âmbito de uma esfera pública digital, e os estudiosos ainda a utilizam como exemplar ideal que serve de comparação com outras esferas públicas empiricamente observadas.

A sétima e última condição de conceitos essencialmente contestados é sua (vii) concorrência progressiva (*progressive competition*). Segundo Rauchfleisch, citando outros pesquisadores, é quase impossível ter uma

---

<sup>75</sup> Cf. HABERMAS, Jürgen. *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy*. Cambridge, MA: MIT Press, 1996.

<sup>76</sup> Cf. CALHOUN, Craig (ed.). *Habermas and the Public Sphere*. Cambridge, MA: MIT Press, (1992).

<sup>77</sup> Suas respectivas contribuições à coleção de ensaios *Habermas and the Public Sphere*, citada na nota anterior, são as seguintes: Benhabib, Seyla. “Models of Public Space: Hannah Arendt, the Liberal Tradition, and Jürgen Habermas” (p. 73-98); Fraser, Nancy. “Rethinking the Public Sphere: A Contribution to the Critique of Actually Existing Democracy” (p. 109-142).

<sup>78</sup> Cf. HABERMAS, Jürgen. “Further Reflections on the Public Sphere”, in: Calhoun, Craig (ed.). *Habermas and the Public Sphere*, op. cit., p. 421-461)

<sup>79</sup> Cf. RAUCHFLEISCH, Adrian. “The public sphere as an essentially contested concept”, op. cit., p. 4.

progressão para um entendimento mútuo, mas é possível avançar no entendimento do conceito com exemplos do mundo real, pois o mundo empírico acaba oferecendo evidências a favor ou contra uma versão proposta por um ou outro pensador, concorrendo progressivamente para sua evolução. E a realidade atual vem oferecendo diversas evidências das alterações na esfera pública.

Como esperado, embora seja um conceito essencialmente contestado, Habermas domina o debate sobre esfera pública, não apenas por seu texto seminal, mas também por obras posteriores, influenciando diversas áreas de pesquisa sobre a esfera pública, especialmente devido ao foco privilegiado na deliberação política. Ele mesmo considera, em raro escrito sobre as raízes biográficas dos dois motivos principais de seu pensamento - o “espaço público” e a “esfera pública política” - que “*a esfera pública, entendida como espaço do trato comunicativo e racional entre as pessoas, é o tema que me persegue a vida toda. De fato*”, prossegue Habermas, “*a tríade constituída pela esfera pública, pelo discurso e pela razão dominou minha vida política e meu trabalho científico*”.<sup>80</sup> Convém notar, como aponta Rauchfleisch, que o pensamento habermasiano é frequentemente citado não só na comunidade da esfera pública tradicional, mas também em comunidades de pesquisas voltadas a tópicos específicos da atualidade, como a Internet<sup>81</sup> e a religião.<sup>82</sup>

Em suma, a pesquisa de Adrian Rauchfleisch demonstra que as publicações com “forte desenvolvimento teórico” sobre o conceito de esfera pública estão no epicentro do debate, ao passo que comunidades com enfoque “disciplinar” (por exemplo, histórico) ou “tópico” (Internet e religião) estão mais afastadas do centro. No entanto, Rauchfleisch argumenta que no futuro a comunidade da “esfera pública online” (digital), objeto do presente trabalho, pode se mover em direção ao epicentro do debate e, ao que tudo indica,

---

<sup>80</sup> Habermas, Jürgen. “Espaço público e esfera política pública. Raízes biográficas de dois motivos de pensamento”, in: *Entre Naturalismo e Religião*. Rio de Janeiro: Tempo Brasileiro, 2007, p. 17-30; aqui: p. 19 (original alemão: *Zwischen Naturalismus und Religion*. Frankfurt: Suhrkamp, 2005).

<sup>81</sup> Cf. Habermas, Jürgen. “Political communication in media society: Does democracy still enjoy an epistemic dimension? The impact of normative theory on empirical research”, *Communication Theory*, 16 (2006): 411-426.

<sup>82</sup> Cf. Habermas, Jürgen. “Religion in the public sphere”, *European Journal of Philosophy*, 14 (2006): 1-25. Esse artigo de Habermas foi retomado na obra *Zwischen Naturalismus und Religion* (“Religião na esfera pública. Pressuposições cognitivas para o ‘uso público da razão’ de cidadãos seculares e religiosos”, in: *Entre Naturalismo e Religião*, op. cit., p. 129-167).



trabalhos sobre esfera pública digital e sobre esfera pública transnacional são os que têm o maior potencial para contribuir com o futuro do desenvolvimento teórico da concepção moderna de esfera pública.

Levando em conta que o objetivo deste subcapítulo é comentar em linhas gerais o conceito de esfera pública, é possível fazer uma (entre muitas aspas) “redescoberta” da esfera pública de Habermas a partir da perspectiva dos espaços públicos virtuais de hoje, sublinhando o contexto histórico e tecnológico do século XVIII e os da atual Sociedade Algorítmica. Em certos aspectos as semelhanças serão apontadas, e, em outros, as diferenças enfatizadas. O resultado poderá ser uma noção geral das principais características da esfera pública digital - na qual a relação da esfera pública com os veículos de comunicação será o enfoque central -, características essas que as atuais (e certamente vindouras) regulações atingirão. É necessário dizer, pois, que o que será afirmado a seguir parte de leituras gerais do momento histórico e da teoria habermasiana da esfera pública. Em momento algum tem-se, aqui, qualquer pretensão em esgotar o assunto, o qual, por sua própria natureza contestatória, é inesgotável. Sequer se pretende abordá-lo com vasta profundidade.

Considerando, então, que este capítulo busca argumentar que a abrangência da esfera pública está diretamente ligada à forma com que se transmite informação, bem como comparar as principais características da esfera pública clássica com as da esfera pública digital, é extremamente necessário reduzir a análise aos aspectos úteis para a futura comparação dos modelos regulatórios dos EUA, da Alemanha e do Brasil. Tendo isto em vista, vale a pena averiguar, em primeiro lugar, sob o aspecto temporal, quem produz informação; em segundo lugar, sob o aspecto tecnológico, quem e como o conteúdo é difundido; e em terceiro lugar, sob quais circunstâncias ele é consumido pelo público alvo.

É de se notar, portanto, que ao menos no que tange à característica disruptiva da era digital quanto a quem produz informação, há mais similitudes do que que imaginamos, pois os parâmetros econômicos que provavelmente impactaram o desenvolvimento tecnológico dos dispositivos de comunicação de hoje são comparáveis aos interesses comerciais e políticos que indiretamente motivaram a emergência da esfera pública burguesa francesa e

inglesa do século XVIII. Neste ponto, Dounia Mahloully (Department of War Studies, King's College, Londres) ressalta, num artigo dedicado justamente a repensar a esfera pública em ambiente digital, que

[e]m primeiro lugar, é importante lembrar que, em certa medida, a esfera pública normativa pode ser considerada como um processo elitista, por meio do qual classes sociais altamente educadas do século XVIII e início do século XIX na França, Alemanha e Grã-Bretanha, ganharam liderança e influência política. De fato, liderar a opinião pública ao se tornar uma figura pública não era apenas o privilégio da elite intelectual, mas também um processo inovador pelo qual a maioria das pessoas dificilmente poderia conceber o resultado cultural e político. De fato, Habermas (1962) demonstra que as ideias de publicidade e a opinião pública se emanaram progressivamente da profusão dos círculos intelectuais compartilhando seu gosto pelas artes e suas visões ideológicas nos salões, cafeterias e círculos internos das elites. Simultaneamente, a evolução da imprensa deu a oportunidade para os líderes intelectuais de serem ouvidos por uma parcela mais ampla da população. Nessa perspectiva, a revolucionária esfera pública do século XVIII certamente democratizou o pensamento racional e transmitiu o legado do Iluminismo. No entanto, Habermas enfatiza o fato de que esse processo de democratização foi parcialmente motivado pela necessidade de as elites burguesas conquistarem poder político e liberalizar os mercados.<sup>83</sup>

Sendo assim, é possível concluir que o exemplar de esfera pública burguesa do século XVIII, analisado por Habermas, também estava sendo alvo de intensas mudanças tecnológicas que impactavam a forma pela qual a informação era disseminada. A segunda semelhança entre a esfera pública normativa e a digital reside, então, na emancipação das subjetividades dos indivíduos: antes uma elite intelectual burguesa aflorando na comunicação de suas subjetividades; agora, diferentemente, qualquer indivíduo conectado. Todavia, no caso da Sociedade Algorítmica, esse fenômeno pode ser

---

<sup>83</sup> Tradução livre do autor do trecho: *"First of all, it is important to recall that, to a certain extent, the normative public sphere can be considered as an elitist process, through which highly educated social classes from the eighteenth century and beginning of the nineteenth century in France, Germany, or Great Britain, gained leadership and political influence. Indeed, leading public opinion by becoming a public figure was not only the privilege of the intellectual elite, but also a groundbreaking process for which most people could hardly understand the cultural and political outcome. Indeed, Habermas (1962) demonstrates that publicity and public opinion themselves progressively emanated from the profusion of intellectual circles sharing their taste for arts and their ideological views in salons, coffee houses, and secret societies. Simultaneously, the evolution of the press gave the opportunity for intellectual leaders to be heard by a broader range of the population. From this perspective, the revolutionising public sphere of the eighteenth century certainly democratised rational thinking and conveyed the legacy of the Enlightenment. However, Habermas emphasises the fact that this democratising process was partly motivated by the need of bourgeois elites to gain political power and liberalise markets"*. (Mahloully, Dounia. "Rethinking the Public Sphere in a Digital Environment: Similarities between the Eighteenth and the Twenty-First Centuries, eSharp, 20 (2013) p. 6.

interpretado pela ascensão do amadorismo nos meios de comunicação e da disseminação de informação, enquanto na esfera pública clássica vimos o protagonismo da elite intelectual da época. Essas novas dinâmicas sociais inevitavelmente afetaram o modo como os cidadãos se envolvem politicamente e o modo como os indivíduos deliberam sobre questões públicas para produzir alguma forma de opinião pública.

É neste contexto de emancipação das subjetividades e da possibilidade de qualquer cidadão se fazer ouvir que também se firma a principal diferença entre os dois contextos históricos: por um lado, hoje em dia, qualquer pessoa com um celular pode se tornar um produtor de conteúdo, disseminador de opinião, ou um influenciador, na melhor acepção do termo. Por outro lado, Dounia Mahlouly lembra que onde a expressão das subjetividades dos indivíduos é descrita por Habermas como um processo benéfico, dado que estimula a crítica entre a população, alguns teóricos argumentam que as individualidades de hoje em dia são expressas de maneira muito superficial, o que paradoxalmente afetaria o processo do pensamento racional, pois os novos meios de comunicação e as novas fronteiras entre espaços privados e públicos permitem que as subjetividades se expressem de maneira menos crítica.<sup>84</sup> Nessa toada, o catedrático Karl-Heinz Ladeur comenta que

não apenas instituições são alteradas no processo de transição para uma nova ordem digital do conhecimento, mas também é possível que surjam novas formas de subjetividade. O novo sujeito poderia ser caracterizado como sujeito relacional, como um sujeito que não encontra mais sua unidade na observação e na internalização do conjunto de regras da sociedade, mas que pode se engajar nas operações com distintas possibilidades, com fragmentações, com figurações fracas de uma autorrelação estável.<sup>85</sup>

Nesse sentido, considerando essas alterações no conteúdo difundido e nos sujeitos pertencentes à esfera pública, do ponto de vista do ativismo online, a ampla expressão das subjetividades pode ser vista como uma vantagem e ao

---

<sup>84</sup> Cf. MAHLOULY, Dounia. "Rethinking the Public Sphere in a Digital Environment: Similarities between the Eighteenth and the Twenty-First Centuries. Idem, *Ibidem*. p. 18.

<sup>85</sup> LADEUR, Karl-Heinz. *Por um novo direito das redes digitais: digitalização como objeto contratual, uso contratual de 'meios sociais', proteção de terceiros contra violações a direitos da personalidade por meio de Cyber Courts*. in Abboud, Georges. Nery Jr., Nelson. e Campos, Ricardo (Coordenadores). *Fake News e Regulação*. Coleção Direito do Estado e Transformação. São Paulo: Thomson Reuters Brasil, 2018. p. 42.

mesmo tempo um rumo desconhecido: a Sociedade Algorítmica atribui a indivíduos conscientes o seu poder político e promove uma cultura democrática, possibilitando que cada cidadão reflita sua própria identidade. No entanto, não está claro até que ponto essa forma de engajamento político é sustentável no contexto de uma democracia, uma vez que o conteúdo produzido - por vezes limitado aos 140 caracteres do Twitter, ou aos 15 segundos de vídeos no Instagram, ou até mesmo ao número reduzido de linhas que outros usuários se darão ao trabalho de ler no Facebook -, muito provavelmente simplificam os discursos e reduzem a robustez da opinião pública, não produzindo, de fato, uma deliberação pública.<sup>86</sup>

Aqui, é fundamental abrir um parêntese para comentar sobre o conceito de democracia deliberativa, pois, de certa forma, o ideal deliberativo tangencia os estudos e a própria concepção da esfera pública. Como sabido, as grandes discussões sobre o tema se deram nos anos da década de 1980 com o crescimento do ativismo em prol dos direitos humanos. Em curtíssimo resumo, a democracia deliberativa consiste na ideia de que a produção legítima de leis e decisões políticas deriva da deliberação pública dos cidadãos e que, dentre várias características, no modelo deliberativo, a democracia possui uma estrutura de condições sociais e institucionais que facilita e promove a discussão entre cidadãos livres e iguais, criando condições favoráveis de participação, associação e expressão, institucionalizando, também, o ideal de justificação política.

De forma sucinta, os principais pensadores sobre o modelo deliberativo de democracia, como Thomas Christiano,<sup>87</sup> Joshua Cohen,<sup>88</sup> Seyla Benhabib<sup>89</sup> e Bernard Manin,<sup>90</sup> argumentam que a deliberação pública tem um valor exclusivamente instrumental no aperfeiçoamento da qualidade das decisões na

---

<sup>86</sup> MAHLOULY, Dounia. "Rethinking the Public Sphere in a Digital Environment: Similarities between the Eighteenth and the Twenty-First Centuries, op. cit. nota 84. p. 10.

<sup>87</sup> CHRISTIANO, Thomas. *A importância da deliberação pública*. In: Melo, Rúrion Soares; Werle, Denílson Luiz (Eds.). *Democracia Deliberativa*. São Paulo: Editora Singular, 2007, p. 81-113.

<sup>88</sup> COHEN, Joshua. *Deliberation and democratic legitimacy*. In: Hamlin, Alan; Pettit, Philip (Eds.). *The Good Polity*. Oxford: Blackwell, 1989, p. 17-34

<sup>89</sup> BENHABIB, Seyla. *Sobre um modelo deliberativo de legitimidade democrática*. In: Melo, Rúrion Soares; Werle, Denílson Luiz (Eds.). *Democracia Deliberativa*. São Paulo: Editora Singular, 2007, p. 47-79.

<sup>90</sup> MANIN, Bernard. *Legitimidade e deliberação política*. In: Melo, Rúrion Soares; Werle, Denílson Luiz (Eds.). *Democracia Deliberativa*. São Paulo: Editora Singular, 2007, p.15-45.

democracia, tendo a igualdade no processo de deliberação um valor intrínseco fundamentado nas exigências de justiça. São apresentados, resumidamente, três tipos diferentes de valores que a deliberação pública pode ter na tomada de decisão, quais sejam: (i) os resultados objetivos da deliberação pública; (ii) o valor intrínseco da deliberação pública; (iii) a condição de justificação política da deliberação pública.

Acerca dos resultados objetivos da deliberação,<sup>91</sup> resumidamente: a deliberação permite aos indivíduos compreenderem melhor os interesses existentes na sociedade, e, por conseguinte, a própria sociedade. Além disso, sob o ponto de vista de que cabeças mais instruídas sobre uma sociedade tomam melhores decisões, temos motivos para pensar que uma democracia que promova a deliberação pública toma decisões melhores. Percebe-se também que mesmo quando a deliberação pública produz mais desacordos do que acordos, do ponto de vista democrático-institucional isso pode ser positivo.

No que tange aos valores intrínsecos da deliberação,<sup>92</sup> os autores frisam que, num ambiente onde as partes deliberam umas com as outras, manifesta-se algum tipo de respeito entre os cidadãos. Somado a isso, o ambiente da deliberação pública melhora as condições cognitivas, habilitando o cidadão a entender melhor o que está em jogo numa decisão política. Nesse contexto, resta claro que a deliberação é valiosa instrumentalmente, sendo um mecanismo para tomar decisões mais bem informadas, pois, considerar uma maior multiplicidade de pontos de vista, em última análise, torna a decisão mais qualificada.

Por fim, a respeito da deliberação como um contexto de justificação política,<sup>93</sup> segundo esta ideia, a deliberação é valiosa por ser uma condição de justificação das instituições sociais, que chegam a certas decisões com ajuda da deliberação pública. Não obstante as múltiplas problematizações sobre o assunto, não se pode negar que uma decisão política que passa pelos procedimentos propostos por um modelo deliberativo de democracia é mais

---

<sup>91</sup> Acerca do tema, ver Seyla Benhabib nota supra e Thomas Christiano nota supra e infra.

<sup>92</sup> Christiano, Thomas. *Procedimento e substância na democracia deliberativa*. In: MELO, Rúrion Soares; WERLE, Denílson Luiz (Eds.). *Democracia Deliberativa*. São Paulo: Editora Singular, 2007, p. 115-144.

<sup>93</sup> COHEN, Joshua. *Deliberation and democratic legitimacy*, op.cit. nota 89.

facilmente justificada do que as decisões que não passem por esses crivos procedimentais. Sendo assim, quanto mais bem deliberada é uma lei, política pública ou, em última instância, uma decisão política, mais chances possui de ser também melhor justificada e aceita.

Partindo, então, da premissa de que a deliberação bem informada é extremamente valiosa para as democracias, neste ponto certamente há diferenças que devem ser especialmente notadas para as finalidades deste trabalho: a Internet e as tecnologias a ela associadas acabaram oferecendo a amadores (cidadãos comuns) a oportunidade de contribuir com seus temas de interesse, confrontar opiniões diferentes e encontrar um público-alvo específico. Amadores adquiriram uma influência que, até há pouco tempo atrás, era privilégio exclusivo de uma elite intelectual burguesa de profissionais e especialistas, que se comunicavam via noticiários e programas de televisão ou colunas de jornais, livros etc. Nota-se hoje em dia que esse reconhecimento social de amadores é particularmente significativo no campo das artes, da cultura popular, da ciência e da política, sendo a difusão de vídeos, memes, piadas, uma característica bastante significativa da comunicação moderna. Tratam-se de formas de comunicação que muito provavelmente tornam o debate público consideravelmente mais raso do que longos textos de jornais e livros.

Agora, especialmente no caso da política, a democratização do debate afeta os parâmetros fundamentais da esfera pública, pois blogueiros e usuários de redes sociais não estão sujeitos a qualquer forma de controle, como padrões ético-jornalísticos e editoriais (*gatekeeping*) mínimos. Neste ponto, é crucial notar que a produção, o conhecimento e a distribuição de informação foram, nas democracias modernas, cada vez mais produzidas dentro de organizações de comunicação, jornais e mídias tradicionais. Acontece que essas empresas tradicionais possuem uma dinâmica de redação, revisão hierárquica de matérias e necessidade legal de nomeação de um redator, ou seja, *standards* profissionais,<sup>94</sup> enquanto os amadores produtores e

---

<sup>94</sup> Sobre este ponto, ver ABOUD, Georges; CAMPOS, Ricardo. *A autorregulação regulada como modelo do Direito Proceduralizado: regulação de redes sociais e proceduralização*. in in Abboud, Georges. Nery Jr., Nelson. e Campos, Ricardo (Coordenadores) . São Paulo: Thomson Reuters Brasil, 2018. p. 31 e 32.

disseminadores de informação da atualidade muitas vezes nem sequer revisam seus textos.

Por esse motivo, o conteúdo produzido pelo amadorismo também tende a ser menos confiável. Se de um lado essa forma online de esfera pública falha em confrontar e discutir questões políticas de maneira crítica e organizada, de outro, ela é composta pela produção de conteúdo informacional por literalmente qualquer cidadão, ampliando a quantidade e a pluralidade de opiniões na esfera pública. Mas, em termos gerais, como isso alterou a esfera pública?

Não se terá aqui a ousadia de responder a essa pergunta, mas aceitando essa indagação como um direcionamento do tema e considerando que este trabalho tem como finalidade debater possíveis regulações às plataformas como forma de combater (ou ao menos diminuir) a desinformação, e que tais regulações necessariamente entram em choque com o princípio da liberdade de expressão, mostra-se necessário adentrar alguns pontos específicos sobre a esfera pública úteis para o nosso trabalho. São eles: primeiro, algumas características da liberdade de expressão na esfera pública que a Internet tornou mais salientes; segundo, os *players* e *stakeholders* envolvidos na dinâmica dos discursos online, os quais serão atingidos direta ou indiretamente pelas regulações.

## 2.2 A ascensão do amadorismo e as novas características da comunicação

Em seu artigo “*Digital Speech and Democratic Culture: A Theory of Freedom of Expression for the Information Society*”,<sup>95</sup> Jack Balkin aponta para cinco características dos tipos de discursos veiculados na Internet que são exemplares para a liberdade de expressão em termos genéricos. Balkin argumenta que, embora não representem facetas novas da rede, a Internet tornou essas características particularmente salientes, remodelando nossa concepção do princípio da liberdade de expressão na esfera pública.

Para o referido autor, a primeira característica é que o discurso na Internet abrange todos os assuntos e modos possíveis de expressão, incluindo o sério, o frívolo, a fofoca, o erótico, o escatológico e o profundo. Portanto, a Internet reflete os gostos populares, a cultura popular e os entusiasmos populares. A segunda característica é que a Internet como um todo é cheia de inovações. O crescimento exponencial da Internet em um período de tempo relativamente curto mostra como pessoas “comuns” podem ser extremamente criativas se tiverem a chance de se expressar. E isto, por sua vez, demonstra o quanto pessoas comuns podem alterar os rumos dos debates na esfera pública quando são autorizadas a serem produtoras ativas, em vez de apenas receptoras passivas de seu mundo cultural, como na época em que os veículos e formas de comunicação eram privilégio de poucos.

Em terceiro lugar, grande parte da fonte dessa criatividade é a capacidade de construir algo novo a partir de algo já existente: os tipos de discursos da Internet, como todo discurso, se apropriam e transformam, imitam e copiam, constroem e se misturam. Isto é particularmente verdadeiro na Internet, pois, pela própria forma na qual o “código” foi concebido, encoraja-se a digitalização, a imitação e a vinculação. Não é difícil de notar que a inovação e a transformação contínuas que vemos nas mídias digitais surgem da capacidade de usar mídias antigas para criar novas. Nessa toada, o discurso na Internet se desenvolve continuamente por meio de *linkagem*, colagem,

---

<sup>95</sup> BALKIN, Jack M., *Digital Speech and Democratic Culture: A Theory of Freedom of Expression for the Information Society* (2004). *Faculty Scholarship Series*. 240. p. 31-32. Disponível em <[https://digitalcommons.law.yale.edu/fss\\_papers/240](https://digitalcommons.law.yale.edu/fss_papers/240)> (acessado em 04.02.19).



anotação, mistura e através do que Balkin chamou de “*routing around and glomming on*”.<sup>96</sup>

Em quarto lugar, a fala na Internet é participativa e interativa. Isso significa dizer que as pessoas não assistem (ou ouvem)<sup>97</sup> mídias na Internet como se fossem as da televisão ou rádio. Em vez disso, elas navegam por elas, programam, publicam, escrevem comentários e adicionam continuamente outros discursos a essas mídias interativas. Assim, a fala na Internet é uma atividade social que necessariamente envolve troca, de maneira que os papéis de leitor, redator, produtor e consumidor de informação são confusos e muitas vezes se misturam.

Finalmente, em quinto lugar, por ser uma atividade social, uma questão de interatividade, de dar e receber; não surpreende que os discursos na Internet criem novas comunidades, culturas e subculturas.

Desta forma, concluiu Balkin que

a natureza populista da liberdade de expressão, sua criatividade, sua interatividade, sua importância para a comunidade e a autoformação sugerem que uma teoria da liberdade de expressão centrada no governo e na deliberação democrática sobre questões públicas é muito limitada. O princípio da liberdade de expressão sempre foi algo maior do que a democracia no sentido restrito de voto e eleições, algo maior até do que a democracia no sentido de deliberação pública sobre questões de interesse público. Se a liberdade de expressão é sobre democracia, trata-se de democracia no sentido mais amplo possível, não apenas no nível da governança, ou no nível da deliberação, mas no nível da cultura. A Internet nos ensina que o princípio da liberdade de expressão é sobre, e sempre foi sobre, a promoção e desenvolvimento de uma cultura democrática.<sup>98</sup>

---

<sup>96</sup> Idem. Ibidem. p. 31.

<sup>97</sup> Um comentário importante sobre a regulação do “novo tipo de rádio”: os Podcasts e sua regulação de conteúdo emergiram como um tema controverso no começo de 2018, depois que a personalidade de direita e o frequente teórico de conspiração, Alex Jones, teve seu podcast “*Infowars*” removido da maioria das plataformas, incluindo Apple, Spotify, Stitcher e RadioPublic. Em meio a uma tempestade na mídia social, as plataformas apressaram-se a banir Jones, algumas vezes poucas horas uma da outra, e muitas vezes sem articular como exatamente o discurso de Jones violava seus termos de uso e política de comunidade. O incidente chamou a atenção para os desafios éticos e logísticos que as plataformas de *podcasting* enfrentam ao equilibrar segurança, diversidade e respeito pelos princípios da liberdade de expressão ao articular o conteúdo que permitem em seus serviços e as dificuldades na implementação dessas políticas de forma consistente. Informação disponível em: <<http://clinic.cyber.harvard.edu/2018/11/29/how-podcast-platforms-respond-to-hate-speech-clinic-releases-new-memo/>> (acessado em 04.02.19).

<sup>98</sup> Tradução livre do autor do trecho “ *The populist nature of freedom of speech, its creativity, its interactivity, its importance for community and self-formation, all suggest that a theory of freedom of speech centered around government and democratic deliberation about public issues is far too limited. The free speech principle has always been about something larger than democracy in the narrow sense of voting and elections, something larger even than*

Em resumo, Jack Balkin está dizendo que, no fundo, o princípio da liberdade de expressão serve para realizar uma cultura democrática; e que uma cultura democrática é muito mais do que um conjunto de procedimentos para resolver disputas. É, ao mesmo tempo, uma característica da vida social e uma forma de organização social que exige um maior compromisso com a igualdade social e política: uma cultura "democrática", então, significa além da democracia como uma forma de autogovernança, democracia como uma forma de vida social na qual barreiras injustas de posição e privilégio são dissolvidas, e em que as pessoas comuns ganham uma palavra maior sobre as instituições e práticas que as moldam e alteram seu futuro; ou seja, uma cultura verdadeiramente democrática é aquela na qual pessoas comuns passam a moldar a esfera pública.<sup>99</sup>

É neste momento que Balkin, ainda em 2004, perpassa o ponto central ao qual se pretende chegar neste trabalho, e o qual se tentou problematizar na primeira parte: a ideia de que as plataformas de comunicação online (ou melhor, o ecossistema informacional *online* como um todo), por serem fundamentais para uma esfera pública que aspira desenvolver uma cultura

---

*democracy in the sense of public deliberation about issues of public concern. If free speech is about democracy, it is about democracy in the widest possible sense, not merely at the level of governance, or at the level of deliberation, but at the level of culture. The Internet teaches us that the free speech principle is about, and always has been about, the promotion and development of a democratic culture.*" in BALKIN, Jack M., *Digital Speech and Democratic Culture: A Theory of Freedom of Expression for the Information Society*, op. cit. nota 96. p. 31-32.

<sup>99</sup> *"What makes a culture democratic, then, is not democratic governance but democratic participation. A democratic culture includes the institutions of representative democracy, but it also exists beyond them, and, indeed undergirds them. A democratic culture is the culture of a democratized society; a democratic culture is a participatory culture. If the purpose of freedom of speech is to realize a democratic culture, why is democratic cultural participation important? First, culture is a source of the self. Human beings are made out of culture. A democratic culture is valuable because it gives ordinary people a fair opportunity to participate in the creation and evolution of the processes of meaning-making that shape them and become part of them; a democratic culture is valuable because it gives ordinary people a say in the progress and development of the cultural forces that in turn produce them. Second, participation in culture has a constitutive or performative value: When people are creative, when they make new things out of old things, when they become producers of their culture, they exercise and perform their freedom and become the sort of people who are free. That freedom is something more than just choosing which cultural products to purchase and consume; the freedom to create is an active engagement with the world. By "culture" I mean the collective processes of meaning-making in a society. The realm of culture, however, is much broader than the concern of the First Amendment or the free speech principle. Armaments and shampoo are part of culture; so too are murder and robbery. And all of these things can affect people's lives and shape who they are."* in Idem. *Ibidem*.

democrática dentro de si mesma, precisam promover a liberdade de expressão de uma tal forma que os cidadãos consigam se envolver em discursos racionais e bem informados sobre a sociedade. Considerando a premissa já exposta de que tal fato não se mostra mais possível em um ecossistema inundado de atores políticos que se utilizam da arquitetura das redes sociais para manipular os debates na esfera pública, causando desinformação generalizada,<sup>100</sup> a conjuntura se mostra quase que um paradoxo em si mesma; a Internet, classicamente enxergada ao mesmo tempo como arena e ferramenta da liberdade de expressão, vem sufocando a cultura democrática que ela mesma visa fomentar, justamente pelo desdobramento da própria liberdade de expressão, sem freios e restrições, dentro do ecossistema informacional da atual Sociedade Algorítmica. Nas palavras de Balkin:

O reino da cultura para fins do princípio da liberdade de expressão é um subconjunto do que os antropólogos estudam como formas de cultura. Refere-se a um conjunto de práticas e mídias sociais historicamente contingentes e produzidas historicamente, que os seres humanos empregam para trocar ideias e compartilhar opiniões. Esses são os métodos, práticas e tecnologias através dos quais o diálogo ocorre e a opinião pública é formada. (...) Não podemos fornecer uma lista exaustiva dessas práticas e mídias precisamente porque as convenções sociais e as tecnologias que as definem estão sempre evoluindo; mesmo assim, parece claro o suficiente que a Internet e outras tecnologias digitais são mídias para a comunicação de ideias, e uma maneira cada vez mais importante para as pessoas expressarem suas ideias e formarem suas opiniões. Elas são mídias centrais - e eu diria cruciais - para a realização de uma cultura democrática. A cultura sempre foi produzida através da participação popular. A tecnologia digital simplesmente torna esse aspecto da vida democrática mais óbvio, mais saliente. A rádio e a televisão são tecnologias de recepção cultural massiva, onde alguns falantes podem atingir públicos de tamanho indefinido. Mas a Internet é uma tecnologia de participação cultural de massa, na qual o público pode dar assim como receber, difundir assim como absorver, criar e contribuir, assim como consumir. A tecnologia digital destaca os valores de uma cultura democrática porque oferece a possibilidade tecnológica de ampla participação cultural.<sup>101</sup>

---

<sup>100</sup> A esta altura do trabalho, talvez seja importante recapitular a ideia geral do primeiro capítulo, que foi a averiguação das consequências negativas trazidas pela desinformação. Sob o ponto de vista democrático, em específico, o problema concentra-se no fato de o atual ecossistema informacional estar operando de maneira disfuncional para a esfera pública democrática, causando polarização, extremismo ideológico e incapacidade dos cidadãos se envolverem em discursos racionais e bem informados sobre a sociedade.

<sup>101</sup> Tradução livre do autor do trecho “ *The realm of culture for purposes of the free speech principle is a subset of what anthropologists study as forms of culture. It refers to a set of historically contingent and historically produced social practices and media that human beings employ to exchange ideas and share opinions. These are the methods, practices, and technologies through which dialogue occurs and public opinion is formed. (...) We cannot give an exhaustive list of these practices and media precisely because the social conventions and*

### 3. UM PANORAMA REGULATÓRIO DAS INTERMEDIÁRIAS DA INTERNET: OLD SCHOOL VERSUS NEW SCHOOL SPEECH REGULATION

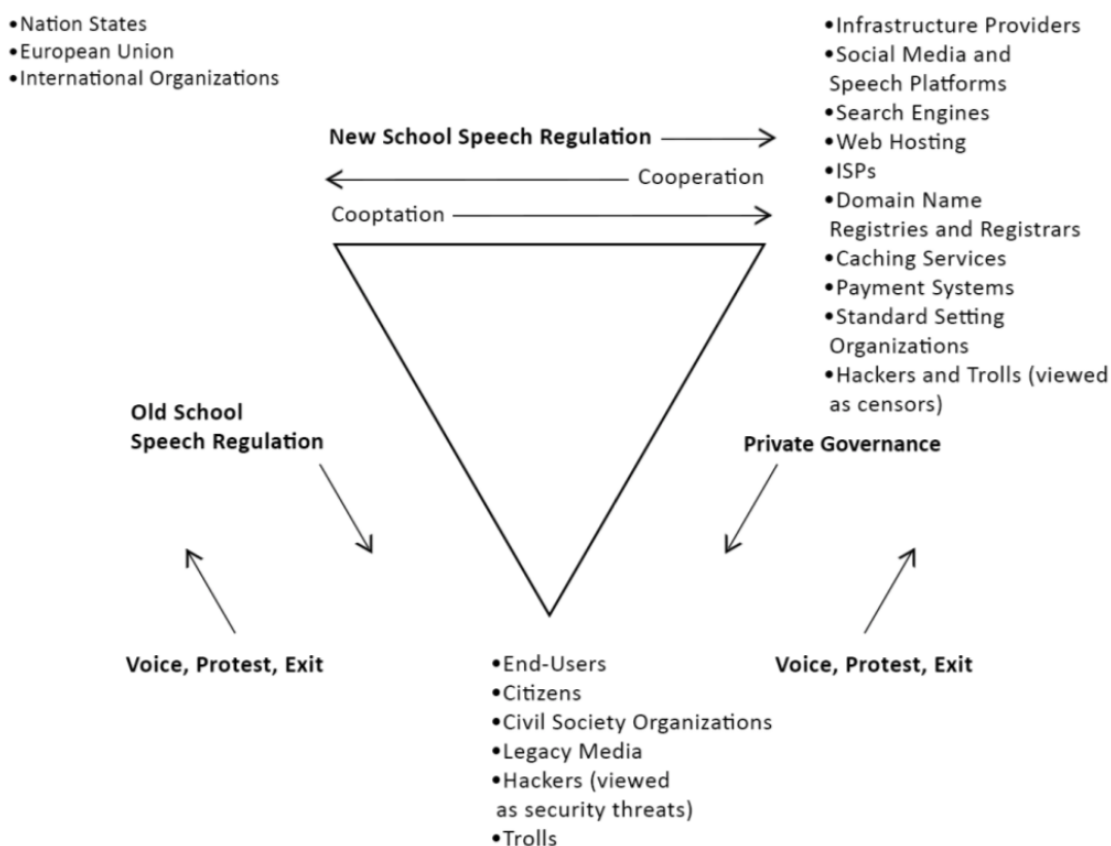
Antes de adentrar o resumo histórico do balizamento das restrições à liberdade de expressão na esfera pública, analisadas adiante por meio da evolução da *public forum doctrine* é necessário delinear o panorama regulatório, ou seja, o cenário regulatório vivenciado pelas empresas de tecnologia que operam plataformas sociais. Neste contexto, no recente artigo “*Free Speech is a Triangle*”<sup>102</sup> de Jack Balkin, fica claro que, para o catedrático, a visão da liberdade de expressão que caracterizou grande parte do século XX é inadequada para proteger o princípio hoje em dia.

Em termos gerais, o professor de Direito Constitucional e Liberdade de Expressão da Faculdade de Direito de Yale explica que o século XX apresentava um modelo dualista (ou diádico) de regulação da fala com dois tipos básicos de atores: os governos ou Estados territoriais, de um lado, e os indivíduos falantes, do outro. Já o modelo regulatório do século XXI é pluralista, com múltiplos *players* e *stakeholders*, de maneira que seria mais fácil enxergar esse modelo regulatório como uma relação triangular. Vejamos.

---

*technologies that define them are always evolving; even so, it seems clear enough that the Internet and other digital technologies are media for the communication of ideas, and an increasingly important way for people to express their ideas and form their opinions.<sup>60</sup> They are central—and I would say crucial—media for the realization of a democratic culture. Culture has always been produced through popular participation. Digital technology simply makes this aspect of democratic life more obvious, more salient. Radio and television are technologies of mass cultural reception, where a few speakers can reach audiences of indefinite size. But the Internet is a technology of mass cultural participation in which audiences can give as well as receive, broadcast as well as absorb, create and contribute as well as consume. Digital technology makes the values of a democratic culture salient to us because it offers the technological possibility of widespread cultural participation.”* in BALKIN, Jack M., *Digital Speech and Democratic Culture: A Theory of Freedom of Expression for the Information Society*, op. cit. nota 96. p. 34-35.

<sup>102</sup> BALKIN, Jack M., *Free Speech is a Triangle*. Columbia Law Review, 2018, Yale Law School, Public Law Research Paper N. 640. Disponível em: <<https://ssrn.com/abstract=3186205>> (acessado em 04.02.19).

FIGURE 1: THE PLURALIST MODEL OF SPEECH REGULATION<sup>10</sup>

No canto superior esquerdo estão os Estados-Nação e a União Europeia, responsáveis pela implementação de regulações. No canto superior direito, estão Intermediários da Internet, em resumo, empresas privadas de infraestrutura de internet, incluindo empresas de mídia social, buscadores, provedores de banda larga e sistemas de pagamento eletrônico. Na terceira quina, há diferentes tipos de falantes, ou seja, sujeitos se expressando na rede.

Balkin, ao diferenciar a “*Old School Speech Regulation*” da “*New School Speech Regulation*” explica que os Estados-Nação continuam a regular os indivíduos e as mídias tradicionais através da regulação tradicional (“*Old School Speech Regulation*”), mas, somado a esta, os Estados também vêm empregando um novo tipo de regulação dos discursos, o qual o professor chama de “*New School Speech Regulation*”, voltada para os Intermediários da Internet. É bom lembrar que os intermediários são considerados pontos de gargalo da comunicação, mostrando-se pontos de contato adequados para

uma regulação jurídica.<sup>103</sup> E por essa razão são setores a serem regulados. O objetivo das novas regulações consiste basicamente, então, em regular a autorregulação das intermediárias, ou seja, levar essas empresas privadas a vigiar, censurar e restringir certos discursos expressos dentro de suas plataformas. Assim sendo, na última aresta estão as empresas Intermediárias da Internet, como o Facebook, que também regulam e governam os indivíduos por meio de técnicas de governança e vigilância privada.

O tema será abordado melhor em momento oportuno, mas, neste ponto é importante destacar a diferença entre regulação, autorregulação e autorregulação regulada, pois esta última vem sendo adotada como modelo de proceduralização do Direito no caso do combate às *fake news* nas redes sociais, e na melhoria da aplicação das leis no mundo digital.

A regulação, segundo Gustavo Binenbojm,<sup>104</sup> envolve uma atividade de interferência no exercício da liberdade privada, ou seja, atuação do Estado no domínio econômico, por meio de coerções ou de induções (poder de polícia) para ordenar as atividades privadas consoante o interesse público. Assim, o Estado regulador define competências e regras de condutas exigíveis dos administrados, conformando o comportamento dos agentes econômicos às razões do Estado (correções de falhas de mercado, de falhas de escolha ou outros fins públicos legítimos, etc), mediante uso de força, se necessário. O autor separa os instrumentos regulatórios em três categorias: (i) regulação por normas de comando e controle; (ii) regulação por normas de indução; (iii) meios alternativos de realização de objetivos regulatórios.

Por seu turno, a “autorregulação”, segundo Karl-Heinz Lauder foi compreendida até agora primariamente como alternativa à regulação estatal de direito público, na qual desenvolveu-se “um direito privado regulador” que, mediante formas de leis de direito privado - de maneira diversa ao modelo de concorrência clássica -, para além de apenas observar as relações de

---

<sup>103</sup> Nesse sentido ver EIFERT, Martin. *A Lei Alemã para a melhoria da aplicação da lei nas redes sociais (NetzDG) e a regulação da plataforma*. in Abboud, Georges. Nery Jr., Nelson. e Campos, Ricardo (Coordenadores). *Fake News e Regulação*. Coleção Direito do Estado e Transformação. São Paulo: Thomson Reuters Brasil, 2018, p. 62.

<sup>104</sup> BINENBOJM, Gustavo. *Poder de polícia, ordenação, regulação: transformações político-jurídicas, econômicas e institucionais do direito administrativo ordenador*. 2. ed. Belo Horizonte: Fórum, 2017, p.143-153.

mercado, eventualmente modificam-nas por meio de intervenções.<sup>105</sup> Se por um lado, a autorregulação tem a vantagem da eficiência pela disposição do conhecimento interno do setor a ser regulado e uma dinâmica de constante revisão e atualização de conceitos a serem regulados,<sup>106</sup> por outro lado, ela se mostra desvantajosa por não necessariamente perseguir interesses e valores públicos.<sup>107</sup>

Já a autorregulação regulada, que na figura acima pode ser representada pela “*New School Speech Regulation*”, busca induzir o setor privado a contribuir para o cumprimento de tarefas públicas. No caso das plataformas, a autorregulação regulada consiste em uma regulação branda imposta pelo Estado às redes sociais para que elas, respeitando a *soft regulation* imposta, se autorregulem e governem suas plataformas por si próprias, sob o balizamento das finalidades públicas. Este modelo opera no liame entre a autorregulação e a regulação pelo Estado, oferecendo outra e nova possibilidade de lidar com as incertezas do mundo tecnológico em constante evolução, inovação e disrupção,<sup>108</sup> ao conciliar as vantagens das duas abordagens: ao mesmo tempo em que almeja interesses e valores públicos, mantém a expertise e dinâmica próprias do setor privado, incorporando conhecimentos de áreas tecnológicas que o Estado não dispõe, ou não dispõe com facilidade e rapidez necessárias para regular.

Para centralizar o debate, é de extrema valia resumir os *players* e *stakeholders* envolvidos, pois, a nível regulatório, eventuais soluções recaem diferentemente sobre cada um deles. Em primeiro lugar estão os consumidores de conteúdo, entendidos como todos os cidadãos que acessam a Internet,

---

<sup>105</sup> LADEUR, Karl-Heinz. *Por um novo direito das redes digitais: digitalização como objeto contratual, uso contratual de ‘meios sociais’, proteção de terceiros contra violações a direitos da personalidade por meio de Cyber Courts*, op. cit. nota 86, p. 45.

<sup>106</sup> José Vicente Santos de Mendonça explica que um dos principais problemas da regulação pública é a informação, pois, o Poder Público não possui domínio epistêmico total sobre a realidade, uma vez que a informação é sempre mais complexa do que a tentativa de instrumentalizá-la. Sobre o tema ver: MENDONÇA, José Vicente Santos de. *Direito Constitucional Econômico: a intervenção do Estado na economia à luz da razão pública e do pragmatismo*. 2. ed. Belo Horizonte: Fórum, 2017, p. 359.

<sup>107</sup> CAMPOS, Ricardo. MARANHÃO, Juliano. *Fake news e autorregulação regulada das redes sociais no Brasil: fundamentos constitucionais*. in Abboud, Georges. Nery Jr., Nelson. e Campos, Ricardo (Coordenadores). *Fake News e Regulação*. Coleção Direito do Estado e Transformação. São Paulo: Thomson Reuters Brasil, 2018, p. 219-220.

<sup>108</sup> Sobre as dificuldades de regular tecnologias disruptivas ver CORTEZ, Nathan. *Regulating Disruptive Innovation*, 29 Berkeley Tech. L.J. (2014).

havendo neste trabalho um especial enfoque para o consumo de conteúdo informacional através das redes sociais. Em segundo lugar estão os criadores de conteúdo (jornalistas profissionais, blogueiros ou “amadores”). Há de se diferenciar, no entanto, os criadores de conteúdo com *standards* profissionais, ou seja, jornais e emissoras de TV, dos criadores de conteúdo sem barreiras ético-profissionais (blogueiros e cidadãos comuns).

Em terceiro lugar, estão os distribuidores de conteúdo. Certamente, neste ponto deve-se problematizar a classificação, pois há um desacordo sobre a melhor forma de distinguir entre diferentes tipos de distribuidores de conteúdo. Alguns favoreceram a distinção entre distribuidores de conteúdo tradicionais (jornais, emissoras) e distribuidores de conteúdo digital (wikis, blogs, plataformas de mídia social, mecanismos de pesquisa, sites de notícias online). Outros favoreceram a divisão com base no fato de um distribuidor de conteúdo ter um processo editorial (jornais, alguns blogs) ou depender da seleção algorítmica para determinar qual conteúdo aparece primeiro (mecanismos de pesquisa ou, por exemplo, as principais plataformas de redes sociais).

Nesse contexto, é importante destacar que existe uma divisão clássica nas camadas da Internet, e este trabalho, por tratar de tema regulatório, está interessado em averiguar com maior profundidade as possibilidades de restrições à liberdade de expressão que são impostas aos “Intermediários da Internet”, sendo fundamental explicar, resumidamente, o que se entende por “intermediário”.

Em relatório de 2010, a OCDE (Organização para a Cooperação e Desenvolvimento Econômico) definiu que Intermediários da Internet “*reúnem ou facilitam transações entre terceiros na Internet. Eles dão acesso, hospedam, transmitem e indexam conteúdo, produtos e serviços originados por terceiros na Internet ou fornecem serviços baseados na Internet a terceiros.*”<sup>109</sup>

Adentrando às minúcias, o relatório atenta para o fato de que a natureza, escopo e escala dos intermediários estão evoluindo rapidamente, sendo crescentes os números e tipos de usuários, informações e serviços prestados.

---

<sup>109</sup> PERSET, Karine /OECD. March 2010. *The Economic and Social Role of Internet Intermediaries*. Paris, Organisation for Economic and Co-operation and Development, p. 9. Disponível em <[www.oecd.org/internet/ieconomy/44949023.pdf](http://www.oecd.org/internet/ieconomy/44949023.pdf)> (acessado em 04.02.19).



Portanto, deve-se notar que, além de serem de natureza muito dinâmica, há diferentes categorias de “Intermediários da Internet”, categorias frequentemente não tão claras, pois, por vezes, um intermediário pode desempenhar mais de um papel de intermediação.<sup>110</sup>

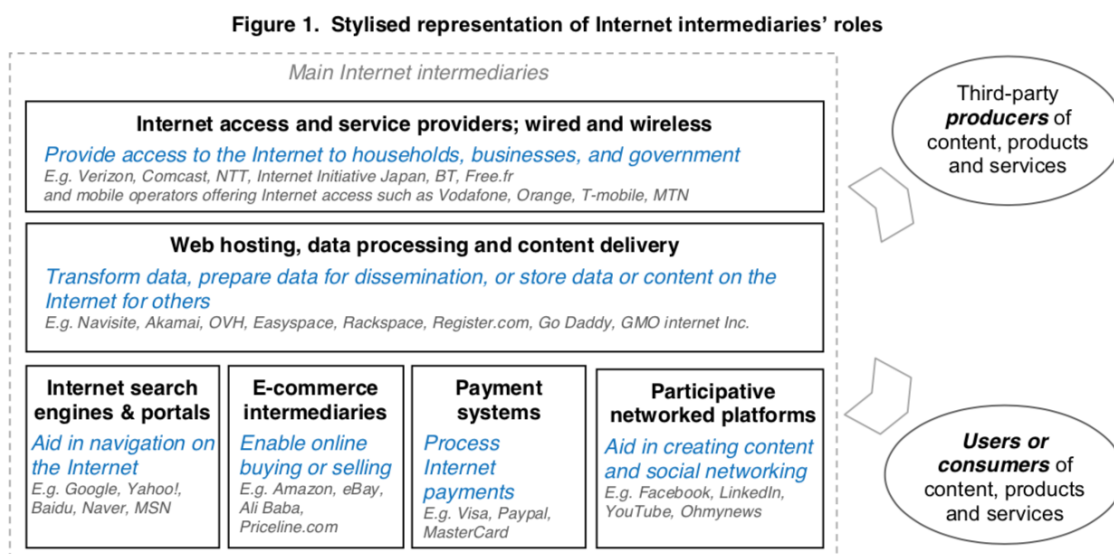
Não obstante exercerem funções híbridas, a OCDE divide os “Intermediários da Internet” em: (i) provedores de acesso à Internet e provedores de serviços (ISPs);<sup>111</sup> (ii) provedores de processamento de dados e hospedagem na Web, incluindo registradores de nomes de domínio; (iii) buscadores e portais da Internet; (iv) intermediários de comércio eletrônico, em que essas plataformas não assumem a propriedade das mercadorias vendidas; (v) sistemas de pagamento pela Internet; e (vi) plataformas de rede participativa, que incluem plataformas de publicação e difusão na Internet que não criam ou detêm o conteúdo que está sendo publicado ou transmitido; ou

---

<sup>110</sup> O relatório também define intermediação nas seguintes palavras: “*Intermediation is the process by which a firm, acting as the agent of an individual or another firm, leverages its middleman position to foster communication with other agents in the marketplace that will lead to transactions and exchanges that create economic and/or social value. The main functions of Internet intermediaries are i) to provide infrastructure; ii) to collect, organise and evaluate dispersed information; iii) to facilitate social communication and information exchange; iv) to aggregate supply and demand; v) to facilitate market processes; vi) to provide trust; and vii) to take into account the needs of both buyers/users and sellers/advertisers. There is sometimes tension between various functions of Internet intermediaries; for example, tension between preserving identity and privacy while personalising products and services in ways that benefit users or between infrastructure provision and usage.*” in Idem. Ibidem. p.6.

<sup>111</sup> Em famoso julgado sobre a responsabilização cível dos provedores de serviços de internet, a Ministra Nancy Andrigui do STJ já havia proferido voto no REsp nº 1.308.830/RS dividindo os provedores de serviços nas seguintes categoriais: “*Os provedores de serviços de internet são aqueles que fornecem serviços ligados ao funcionamento dessa rede mundial de computadores, ou por meio dela. Trata-se de gênero do qual são espécies as demais categorias, como: (i) provedores de backbone (espinha dorsal), que detêm estrutura de rede capaz de processar grandes volumes de informação. São os responsáveis pela conectividade da internet, oferecendo sua infraestrutura a terceiros, que repassam aos usuários finais acesso à rede; (ii) provedores de acesso, que adquirem a infraestrutura dos provedores backbone e revendem aos usuários finais, possibilitando a esses conexão com a internet; (iii) provedores de hospedagem, que armazenam dados de terceiros, conferindo-lhes acesso remoto; (iv) provedores de informação, que produzem as informações divulgadas na internet; e (v) provedores de conteúdo, que disponibilizam na rede as informações criadas ou desenvolvidas pelos provedores de informação. É frequente que provedores ofereçam mais de uma modalidade de serviço de internet; daí a confusão entre essas diversas modalidades. Entretanto, a diferença conceitual subsiste e é indispensável à correta imputação da responsabilidade inerente a cada serviço prestado. Na hipótese específica do ORKUT, rede social virtual na qual foram veiculadas as informações tidas por ofensivas, verifica-se que a GOOGLE atua como provedora de conteúdo, pois o site disponibiliza informações, opiniões e comentários de seus usuários. Esses usuários criam páginas pessoais (perfis), por meio das quais se relacionam com outros usuários e integram grupos (comunidades), igualmente criados por usuários, nos quais se realizam debates e troca de informações sobre interesses comuns.*”

seja, as redes sociais. A divisão é melhor sintetizada por essa tabela, constante no relatório da OCDE:



É bom especificar que quando estamos falando de regulação no contexto da desinformação, o escopo das restrições incidirão predominantemente em dois tipos de Intermediários, representados nos dois vértices inferiores da figura acima: buscadores e portais da Internet,<sup>112</sup> como, por exemplo, no caso das incidências do Direito ao Esquecimento na Internet,<sup>113</sup> e nas plataformas de rede participativa, as redes sociais.

Embora os debates se tangenciem, neste trabalho há nitidamente um enfoque nas regulações impostas ao último tipo de Intermediário, as

<sup>112</sup> No relatório, a OCDE define que “*Internet search engines and portals operate websites that use a search engine to generate and maintain extensive databases of Internet addresses and content in an easily searchable format. Content may consist of web pages, images or other types of digital files. Search engines index information and content in an automated fashion, based on sophisticated algorithms. Web search portals often provide additional Internet services, such as e-mail, connections to other websites, auctions, news, and other limited content. It should be noted that many portals do not rely on automated search engines alone, but also include human editors whose function is similar to that of a magazine editor. Search engines and portals generally provide free services to their users even though these services involve significant investment in technical development and infrastructure to meet simultaneous demand of a growing number of users. Investments and operating costs are most often funded through advertising. For example, Google, Naver in Korea and Baidu in China, use auction-based advertising programs that let advertisers deliver ads targeted to search queries or web content across the search-engines’ sites and through affiliated third party websites. Advertisers are increasingly charged per user that clicks on the ad versus per user that sees the ad. Revenue-sharing mechanisms with affiliated websites are often used.*”, op. cit nota 110. p. 12.

<sup>113</sup> Sobre o tema ver a obra BRANCO, Sérgio. *Memória e esquecimento na internet*. Porto Alegre: Arquipélago Editorial, 2017.

plataformas de interação social, pois são nas redes sociais como o Facebook, Twitter, Instagram e o Reddit, que a desinformação vem operando em maior escala (muito embora o aplicativo “WhatsApp” tenha papel especial no contexto brasileiro de desinformação,<sup>114</sup> não se pode classificá-lo cegamente como plataforma interativa, pois trata-se de aplicativo de envio de mensagens, com natureza diferente).<sup>115</sup> Além disso, a outra razão - que será demonstrada no próximo subtópico -, consiste no fato de que estas plataformas podem ser entendidas atualmente como componentes da esfera pública, sendo o “locus” no qual as regulações buscam atingir, pois são consideradas gargalos da comunicação, mais indicadas para regulações jurídicas. Portanto, a OCDE explica que as plataformas de rede participativa

facilitam a comunicação social e a troca de informações. São serviços baseados em novas tecnologias, como a Web, mensagens instantâneas ou tecnologias móveis que permitem aos usuários contribuir para o desenvolvimento, classificação, colaboração e distribuição de conteúdo da Internet e desenvolvimento e personalização de aplicativos da Internet ou para a realização de redes sociais. Destina-se a incluir sites de redes sociais, sites de conteúdo de vídeo, sites de jogos online e mundos virtuais.<sup>116</sup>

Vejamos a tabela exemplificativa contida no relatório da OCDE:

---

<sup>114</sup> Como relatado na primeira parte, o WhatsApp vem investindo em melhorias para evitar a desinformação em massa. Essas atualizações são implementadas através de novas regras no aplicativo, inseridas pela alteração de ferramentas dentro do sistema. Alguns dos exemplos de mudança na plataforma são: (i) a restrição a grupos para até 250 membros; (ii) a criação da advertência “encaminhada” quando mensagens são compartilhadas de uma janela para outra (sinalizando a possibilidade de ser uma corrente tentando propagar informação em cascata); (iii) restrições na quantidade de usuários em listas de transmissão e compartilhamento de mensagens. Não está errado quem diz que a implementação dessas alterações é uma forma de autogovernança e, portanto, nada mais do que uma autorregulação.

<sup>115</sup> Como comentado na primeira parte, o uso do WhatsApp como canal de distribuição de *fake news* tem especial relevância. Sobre o tema ver recente estudo do ITS-RIO, que pode ser acessado em: <<https://feed.itsrio.org/poder-computacional-automa%C3%A7%C3%A3o-no-uso-do-whatsapp-nas-elei%C3%A7%C3%B5es-e969746d231f>> (acessado em 04.02.19).

<sup>116</sup> Tradução livre do autor do trecho “*Participative networked platforms facilitate social communication and information exchange. They are services based on new technologies such as the web, instant messaging, or mobile technologies that enable users to contribute to the developing, rating, collaborating and distributing Internet content and developing and customising Internet applications, or to conduct social networking. This category is intended to include social networking sites, video content sites, online gaming websites and virtual worlds. Table 2 provides an overview of well-known participative networked platforms.*” in op cit. nota 110. p. 14.

Table 2. Participative networked platforms

Type of Platform	Examples
Blogs	Blogs such as BoingBoing, Engadget, Ohmy News; Blogs on sites such as LiveJournal; Windows Live Spaces; Cyworld; Skyrock
Wikis and other text-based collaboration formats	Wikipedia, Wiktionary; Sites providing wikis such as PBWiki, Google Docs
Instant messaging	Skype, Trillian, Windows Live Messenger
Mobile	Mobile versions of social networking sites and applications such as Facebook
Sites allowing feedback on written works	FanFiction.Net, SocialText, Amazon
Group-based aggregation	Sites where users contribute links and rate them such as Digg, reddit Sites where users post tagged bookmarks such as del.icio.us
Photosharing sites	Kodak Gallery, Flickr
Podcasting	iTunes, FeedBurner (Google), WinAmp, @Podder
Social network sites	MySpace, Mixi, Facebook, Twitter, Bebo, Orkut, Cyworld, Imeem, ASmallWorld
Virtual worlds <sup>23</sup>	Second Life, Active Worlds, Entropia Universe, Dotsoul Cyberpark
Online computer games	World of Warcraft, Tomb Raider, Lineage Ultima Online, Sims Online, Club Pogo <sup>24</sup>
Video content or filesharing sites	YouTube, DailyMotion, GyaO, Crackle

Source: Building on OECD, *Information Technology Outlook 2008, Chapter 5 - Digital Content and Convergence in Transition*.

Resumidamente, essas plataformas interativas geralmente são baseadas em modelos de comunidade, regulamentados por “termos de uso” que o usuário precisa aceitar e respeitar para ter acesso ao serviço. Embora os modelos de negócios das redes sociais ainda estejam em fluxo dinâmico de mudança, e considerando que os usuários investem uma alta quantidade de tempo nessas plataformas, essas empresas baseiam sua receita nas mais variadas formas de captação de recursos pela exploração da principal *commodity* das plataformas: a atenção dos usuários. Na prática, angariam recursos pela venda de produtos e serviços auxiliares, doações voluntárias, publicidade e assinaturas de serviços chamados *premium*.

É importante notar que a maioria das definições de intermediários exclui explicitamente os produtores de conteúdo, como aponta publicação de 2014 da UNESCO sobre o papel das Intermediárias da Internet.<sup>117</sup> Em outras palavras, no contexto do ecossistema informacional, editores e outras mídias que criam e disseminam conteúdo não seriam intermediários, a exemplo de sites de notícias que publicam artigos escritos e editados por funcionários ou contribuidores convidados, ou um serviço de assinatura de vídeo digital que contrata ou convida pessoas a produzirem vídeos, disseminando-os para assinantes. É bom lembrar, no entanto, que muitas dessas entidades oferecem serviços híbridos e acabam se encaixando como intermediários de uma forma ou de outra, de maneira que a distinção acaba perdendo um pouco de sentido.

<sup>117</sup> MACKINNON, Rebecca. HICKOK, Elonnai. BAR, Allon. Lim, Hae-in. *Fostering Freedom Online: The role of Internet Intermediaries*. Series on Internet Freedom. Paris: United Nations Educational, Scientific and Cultural Organization (UNESCO), 2014, p. 20.

Esse é particularmente o caso das organizações de mídia tradicionais que também têm presença nas redes sociais e fornecem ou facilitam o compartilhamento de conteúdo gerado pelo usuário, constituindo, então, intermediários híbridos. Será sempre crucial, todavia, avaliar em que medida um determinado serviço de mídia social, por exemplo, é principalmente um intermediário ou um intermediário que também opera uma função de mídia.

Para o escopo deste trabalho, nota-se que os Intermediários da Internet estudados aqui exigem que os usuários concordem com os "termos de serviço" (ou "termos de uso") antes de poderem desfrutar do serviço.<sup>118</sup> Às vezes, esses termos podem ser utilizados para restringir o discurso dos usuários que, na verdade, não são proibidos pela lei em algumas jurisdições. Por exemplo: o Facebook e o Twitter proíbem a nudez e várias formas de discurso de ódio em suas plataformas, embora a maioria desses conteúdos não seja ilegal em diversos ordenamentos jurídicos.<sup>119</sup> Embora a aplicação de tais termos possa se assemelhar a uma função editorial, a base para a execução dos termos de serviço nos EUA e na Europa (em que os intermediários da Internet surgiram) - e também no Brasil - não é derivada da lei de mídia ou imprensa, mas do direito contratual e comercial vinculado à aplicação dos termos de uso, pelo qual as empresas têm o direito de exigir que os usuários cumpram os termos de uso em troca do serviço.

Voltando ao panorama regulatório, em artigo que buscou justamente diferenciar o antigo modelo regulatório do novo, Jack Balkin ensina que a "*Old School Speech Regulation*" trata de uma regulação aos falantes, aos espaços e às tecnologias tradicionais (Pré-Digitais) de publicação; portanto, são restrições que englobam espaços públicos, como agências de correios, editoras, filmes, telegrafia, telefonia, radiodifusão e televisão. Em suma, os Estados-Nação -

---

<sup>118</sup> Ver, por exemplo, Padrões da Comunidade, Facebook, <<http://www.facebook.com/communitystandards>>; Termos de Uso, Instagram, <<https://www.facebook.com/help/instagram/478745558852511>>; Política de Conteúdo Reddit, Reddit, <<http://www.reddit.com/help/contentpolicy>>; Suporte Snapchat, Guia da Comunidade, Snapchat, <<http://support.snapchat.com/pt-BR/a/guidelines>>; Twitter Help Ctr., As Regras do Twitter, Twitter, <<https://help.twitter.com/pt/rules-and-policies/twitter-rules>>; Diretrizes da comunidade, Youtube, <<http://www.youtube.com/yt/policyandsafety/communityguidelines.html>> (acessados em 04.02.19).

<sup>119</sup> Sobre o tema, um debate que se tornou muito caloroso nas redes sociais, especialmente no Instagram, com a ascensão do feminismo, é a proibição de veiculação de imagens contendo mamilos femininos. Por vezes, postagens são retiradas do ar por supostamente violarem as regras da comunidade.

seja por regulações diretas ou pela construção de decisões jurisprudenciais - acabam exercendo o poder regulatório por meio de controle monopolista sobre as tecnologias de transmissão, ou indiretamente através do licenciamento, restringindo quem pode usar essas tecnologias, e quais são os limites aos discursos nesses espaços.

Já a “*New School Speech Regulation*”, para Balkin, trata de uma regulação de controle sobre os intermediários e plataformas digitais, e de censura colateral. Na atualidade, como explicado acima, as plataformas digitais e os intermediários da Internet acabam desempenhando funções híbridas, por exemplo, se unindo a emissoras de televisão, cabo e rádio, dificultando o uso das regulações naturais ao modelo antigo para definir balizamentos à fala. Soma-se a isso o fato de que, no contexto digital, muitos usuários são anônimos, utilizam pseudônimos ou estão localizados no exterior, fora do alcance territorial da jurisdição do Estado e das suas restrições. Portanto, para o catedrático, neste cenário, os governos precisam se voltar para outros modelos regulatórios: em vez de (ou além de) regular os indivíduos se expressando, os governos devem visar os intermediários da Internet, ou melhor, as plataformas nas quais os discursos são expressos - os gargalos comunicacionais.

Apenas para que se entenda a construção das proteções legais concedidas às redes sociais analisadas logo a seguir, é importante resumir que os problemas envolvendo censura colateral, ou seja, restrições impróprias ao princípio da liberdade de expressão, normalmente ocorrem em países com modelos de responsabilização por mera notificação extrajudicial.<sup>120</sup> Os modelos

---

<sup>120</sup> Um debate interessante sobre o tema será travado pelo Supremo Tribunal Federal em breve, pois a constitucionalidade do artigo 19 do Marco Civil da Internet (Lei Federal nº 12.965/14) foi questionada no Recurso Extraordinário nº 1.037.396 interposto pelo Facebook Brasil. O relator do caso, ministro Dias Toffoli, reconheceu a sua repercussão geral, abrindo caminho para que se altere o modelo de responsabilização adotado pelo legislador brasileiro no referido artigo. O dispositivo pacificou uma jurisprudência oscilante nos Tribunais, inserindo o modelo de responsabilização por ordem judicial no Brasil. Neste modelo, as plataformas de internet (os “intermediários” ou “provedores de aplicações”) são passíveis de responsabilização por conteúdos publicados por seus usuários apenas quando, após serem notificados de decisão judicial específica determinando a sua remoção, não tomarem providências para tornar indisponível o conteúdo apontado como ilícito pelo Poder Judiciário. A lei estabeleceu exceções para casos envolvendo direito autoral (art. 19, §2º) ou a divulgação não consensual de imagens íntimas (art. 21). Sobre o tema ver SOUZA, Carlos Affonso. LEMOS, Ronaldo. *Marco civil da internet: construção e aplicação*. Juiz de Fora: Editar Editora Associada Ltda, 2016. Disponível em <[https://itsrio.org/wp-content/uploads/2017/02/marco\\_civil\\_construcao\\_aplicacao.pdf](https://itsrio.org/wp-content/uploads/2017/02/marco_civil_construcao_aplicacao.pdf)> (acessado em 04.02.19).

regulatórios adotados nestes casos são lastreados em práticas que privilegiam o controle da circulação de conteúdos em detrimento da liberdade de expressão e do acesso irrestrito à informação, pois o Estado mantém uma parte privada (a rede social) responsável pela fala de outra parte privada (usuário da plataforma). Nesse modelo, os intermediários de Internet são passíveis de responsabilização a partir do momento em que são comunicados extrajudicialmente, pelo interessado, da existência de conteúdo alegadamente indesejado. Na hipótese do conteúdo violar direito de terceiro, a vítima poderia notificar diretamente o provedor do serviço, que deve, então, bloquear prontamente o acesso ao conteúdo ou apagá-lo. Se, após a notificação, o intermediário não adotar as medidas necessárias para impedir o acesso ao referido conteúdo, ele poderá ser responsabilizado.

No entanto, o que de fato acontece na prática do mercado é que, como a rede social detém o poder de bloquear, censurar ou de outra forma controlar os discursos dos usuários, com base nos seus termos de uso ou políticas de serviço, a plataforma acaba optando por bloquear o discurso de seus usuários ou retirar o seu acesso à rede, visando se resguardar de qualquer tipo de responsabilidade cível ou criminal. E, de fato, neste cenário de responsabilização por mera notificação extrajudicial, comum a países com vestígios autoritários,<sup>121</sup> as redes sociais não têm incentivos para manter o conteúdo no ar e, ao optarem pela cautela, podem acabar restringindo até mesmo discursos totalmente protegidos pela liberdade de expressão, para evitar qualquer chance de responsabilização.<sup>122</sup> E pior, podem fazê-lo sem a devida transparência e controle externo.

---

<sup>121</sup> Países como China, Rússia, Irã, Ruanda e Venezuela adotam o modelo de responsabilização por notificação extrajudicial. Sobre o tema ver manifestação de Amicus Curiae da Associação Internetlab de pesquisa em direito e tecnologia no Recurso Extraordinário 1.037.396/SP, disponível em <<http://www.internetlab.org.br/wp-content/uploads/2018/12/Amicus-Curiae-InternetLab-RE-1037396-assinado.pdf>>. Ver também Freedom House. *Freedom on the Net 2017: Manipulating Social Media to Undermine Democracy*, op cit. nota 37.

<sup>122</sup> Sobre o tema, ver coleção de estudos acadêmicos documentando a remoção de conteúdo legal por intermediários como consequência de sistemas jurídicos que adotam o modelo de responsabilização por mera notificação extrajudicial em várias jurisdições em KELLER, Daphne. *Empirical Evidence of "Over-Removal" By Internet Companies Under Intermediary Liability Laws*, The Center for Internet and Society at Stanford Law School, (2017). Disponível em <<http://cyberlaw.stanford.edu/blog/2015/10/empirical-evidence-over-removal-internet-companies-under-intermediary-liability-laws>> (acessado em 04.02.19).

No modelo regulatório norte-americano, os intermediários gozam de vários níveis diferentes de proteção, dependendo do conteúdo envolvido. Por exemplo, a Seção §230<sup>123</sup> da Lei de Telecomunicações dos Estados Unidos (*Telecommunications Act of 1996*) mantém provedores de serviços interativos isentos de responsabilização por conteúdo ofensivo publicado em sua plataforma, não se aplicando essa isenção à responsabilização baseada em violação de propriedade intelectual. Já a Seção §512<sup>124</sup> da Lei de Direitos Autorais (*Digital Millennium Copyright Act of 1998*) oferece uma imunidade de responsabilização para provedores de serviços de Internet, operadoras de *backbone* e semelhantes quando conteúdo potencialmente infrator flui através de sua infraestrutura sem o conhecimento da empresa. Dentro da lógica dos Direitos Autorais, a Seção §512 também cria um procedimento elaborado de notificação e remoção (*notice and takedown*)<sup>125</sup> para intermediários (como o YouTube) que hospedam o conteúdo de outras pessoas. Nesse sentido, Balkin argumenta que

as imunidades da Seção § 230 e, em menor medida, as imunidades do § 512 têm estado entre as mais importantes proteções da liberdade de expressão nos Estados Unidos na era digital. Elas tornaram possível o desenvolvimento de uma ampla gama de sistemas de telecomunicações, motores de busca, plataformas e serviços em nuvem sem medo de comprometer a responsabilidade. [...] no entanto, resultaram de atos legislativos em vez de decisões da Suprema Corte. E nem todos os países têm regras de proteção de discursos em responsabilidade por intermediação.<sup>126</sup>

De fato, essas proteções foram necessárias para que essas empresas pudessem desenvolver seu serviço sem serem alvo de ações judiciais que visavam à responsabilização das plataformas por conteúdo difundido por

---

<sup>123</sup> §230(c)(1) (“No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider.”)

<sup>124</sup> § 512(a) (providing immunity for “transitory digital network communications”); id. § 512(b) (providing immunity for temporary caching).

<sup>125</sup> § 512(g) (describing notice-and-takedown procedure for service providers).

<sup>126</sup> Tradução livre do autor do trecho “Section 230 immunity and, to a lesser extent, § 512 safe harbors have been among the most important protections of free expression in the United States in the digital age. They have made possible the development of a wide range of telecommunications systems, search engines, platforms, and cloud services without fear of crippling liability. [...] Both the § 230 immunity and the § 512 safe harbors, however, resulted from legislative acts rather than Supreme Court decisions. And not all countries have speech-protective rules of intermediary liability.” in BALKIN, Jack M., “Old-School/New-School Speech Regulation” (2014). Faculty Scholarship Series. 4877. Disponível em <[https://digitalcommons.law.yale.edu/fss\\_papers/4877](https://digitalcommons.law.yale.edu/fss_papers/4877)>(acessado em 04.02.19).



usuários, algo que, sem dúvida, favoreceu o desenvolvimento desse mercado tecnológico. No mesmo sentido, Nabiha Syed, em seu artigo “*Real Talk About Fake News: Towards a Better Theory for Platform Governance*”,<sup>127</sup> cujo objetivo central foi resumir as principais teorias sobre a Primeira Emenda, ou seja, as construções jurisprudenciais sobre liberdade de expressão nas quais as plataformas se ancoraram ao longo da história, também argumenta que essas imunidades tinham funções claras:

as plataformas de comunicação privada, como o Facebook, o Twitter, o Reddit e o YouTube, não são tão claramente vinculadas pela doutrina da Primeira Emenda quanto seus predecessores. Pelo contrário, essas plataformas gozam de ampla imunidade em relação à responsabilidade, com base nas mensagens, fotografias e vídeos gerados pelos usuários que preenchem suas páginas: há muito tempo, a Seção § 230 [...] oferece amplo ancoradouro para elas construírem suas plataformas como bem entenderem. [...] O objetivo dessa concessão de imunidade foi tanto encorajar as plataformas a serem ‘bons samaritanos’ quanto ter um papel ativo na remoção de conteúdos ofensivos, mas também evitar problemas de liberdade de expressão de censura colateral.<sup>128</sup>

No caso do Brasil, o artigo 19<sup>129</sup> do Marco Civil da Internet (Lei nº 12.965/14) dispõe que, com o intuito de assegurar a liberdade de expressão e

<sup>127</sup> NABIHA Syed, *Real Talk About Fake News: Towards a Better Theory for Platform Governance*, 127 Yale L.J. F. 337 (2017). Disponível em

<<http://www.yalelawjournal.org/forum/real-talk-about-fake-news>> (acessado em 04.02.19).

<sup>128</sup> Tradução livre do autor do trecho “*private communications platforms like Facebook, Twitter, Reddit, and YouTube are not as clearly bound by First Amendment doctrine as their predecessors might have been. To the contrary, these platforms enjoy broad immunity from liability based on the user-generated messages, photographs, and videos that populate their pages: Section 230 of the Communications Decency Act has long given them wide berth to construct their platforms as they please. [...] The purpose of this grant of immunity was both to encourage platforms to be “Good Samaritans” and take an active role in removing offensive content, but also to avoid free speech problems of collateral censorship.*” in Idem, *Ibidem*. p. 339.

<sup>129</sup> Art. 19 da Lei nº 12.965/2014 - Com o intuito de assegurar a liberdade de expressão e impedir a censura, o provedor de aplicações de internet somente poderá ser responsabilizado civilmente por danos decorrentes de conteúdo gerado por terceiros se, após ordem judicial específica, não tomar as providências para, no âmbito e nos limites técnicos do seu serviço e dentro do prazo assinalado, tornar indisponível o conteúdo apontado como infringente, ressalvadas as disposições legais em contrário. § 1º A ordem judicial de que trata o caput deverá conter, sob pena de nulidade, identificação clara e específica do conteúdo apontado como infringente, que permita a localização inequívoca do material. § 2º A aplicação do disposto neste artigo para infrações a direitos de autor ou a direitos conexos depende de previsão legal específica, que deverá respeitar a liberdade de expressão e demais garantias previstas no art. 5º da Constituição Federal. § 3º As causas que versem sobre ressarcimento por danos decorrentes de conteúdos disponibilizados na internet relacionados à honra, à reputação ou a direitos de personalidade, bem como sobre a indisponibilização desses conteúdos por provedores de aplicações de internet, poderão ser apresentadas perante os juizados especiais. § 4º O juiz, inclusive no procedimento previsto no § 3º, poderá antecipar,

impedir a censura, o provedor de aplicações de Internet somente poderá ser responsabilizado civilmente, por danos decorrentes de conteúdo gerado por terceiros, se, após ordem judicial específica, não tomar as providências para, no âmbito e nos limites técnicos do seu serviço e dentro do prazo assinalado, tornar indisponível o conteúdo apontado como infringente, ressalvadas as disposições legais em contrário. Com isso, o Marco Civil da Internet se aproxima do modelo norte-americano e europeu, estabelecendo que a responsabilidade civil do provedor de aplicações de Internet será subjetiva, e derivará apenas em casos de não cumprimento da ordem judicial que determinou a remoção de certo conteúdo.<sup>130</sup>

Considerando o fato de que em momento oportuno também comentaremos sobre a recente norma alemã para melhoria da aplicação da lei nas plataformas (NetzDG) e que o objetivo deste subtópico é traçar o cenário regulatório, cumpre resumir também o modelo regulatório alemão.

O professor catedrático de Direito Público e Administrativo da Humboldt-Universität de Berlim, Martin Eifert, explica que na Alemanha também se construiu ao longo do tempo isenções específicas de responsabilização das plataformas,<sup>131</sup> como, por exemplo, as respaldadas nos §§ 8-11 da Lei dos Serviços à Distância - Teleserviços (*Teledienstegesetz*), no artigo 14 da Diretiva E-Commerce da União Europeia, e os §§7-10 da Lei de Telemídia alemã (*Telemediengesetz*).

O Diretor do “Humboldt Institute for Internet and Society” e Professor da Universidade de Hamburgo, Wolfgang Schulz, em seu artigo “*The Case of the German NetzDG*” também lembra que

Nos termos dos artigos 12.º a 14.º da diretiva 2000/31 / CE (diretiva relativa ao comércio eletrônico), existe um limite para a responsabilidade dos fornecedores de serviços pelo conteúdo de terceiros. A diretiva relativa ao comércio eletrônico (artigo 14) levou ao desenvolvimento de procedimentos de retirada, mas não os regula em pormenor. Embora existam diferenças de escopo e conceito, isso

---

total ou parcialmente, os efeitos da tutela pretendida no pedido inicial, existindo prova inequívoca do fato e considerado o interesse da coletividade na disponibilização do conteúdo na internet, desde que presentes os requisitos de verossimilhança da alegação do autor e de fundado receio de dano irreparável ou de difícil reparação.

<sup>130</sup> Souza, Carlos Affonso. LEMOS, Ronaldo. *Marco civil da internet: construção e aplicação*, op.cit. nota 121.

<sup>131</sup> EIFERT, Martin. *A Lei Alemã para a melhoria da aplicação da lei nas redes sociais (NetzDG) e a regulação da plataforma*, op.cit. nota 104. p. 61.

pode ser visto como o equivalente europeu da seção 230 do US Communications Decency Act.<sup>132</sup>

É de se notar, portanto, que igualmente ao caso norte-americano, a Alemanha, seguindo a diretiva da União Europeia como um todo, também desenvolveu a sua “*Old School Speech Regulation*” fundada no objetivo de alcançar a superação de obstáculos para o livre desenvolvimento do mercado no âmbito dos novos serviços de comunicação, bem como a garantia das condições conjunturais econômicas para a oferta e para o uso dessas novas tecnologias, pois, sem essas isenções de responsabilização por conteúdo - divulgado ou que simplesmente atravesse a infraestrutura de comunicação -, seria economicamente impossível que as empresas intermediárias da internet, e seus serviços, se desenvolvessem.

---

<sup>132</sup> Tradução livre do autor para o trecho “*Under Article 12 to 14 of Directive 2000/31/EC (e-Commerce Directive), there is a limit to the liability of service providers for third party content. The e-Commerce Directive (Article 14) has led to the development of take-down procedures, but does not regulate them in detail. Even though there are differences in scope and concept this can be seen as the European equivalent of section 230 of the US Communications Decency Act.*” in SCHULZ, Wolfgang, *Regulating Intermediaries to Protect Privacy Online – The Case of the German NetzDG* (July 19, 2018). Marion Albers and Ingo Sarlet (editors), *Personality and Data Protection Rights on the Internet*. Disponível em <<https://www.hiig.de/wp-content/uploads/2018/07/SSRN-id3216572.pdf>> (acessado em 04.02.19).

### 3.1 Quais danos buscamos evitar e quais os principais desafios ao regular tecnologias como as plataformas interativas?

Posto, assim, em termos gerais os cenários regulatórios norte-americano e alemão, os quais, em resumo, buscaram eximir as plataformas de responsabilização por conteúdo divulgado por terceiro, é necessário condensar o debate explicando que as plataformas podem ser chamadas a agir em situações de ofensa à lei - momento no qual serão obrigadas a fazer uma ponderação entre as restrições legais ou jurisprudenciais ao direito de expressão e a própria liberdade de expressão. Em síntese, essas plataformas podem ser demandadas a aplicar um juízo de averiguação de legalidade dos discursos, ou seja, uma forma de aplicação do Direito em relações privadas.

Em suma, as plataformas podem ser instadas a averiguar o conteúdo da fala via notificação - judicial ou extrajudicial -,<sup>133</sup> o que as obriga a manterem estruturas internas de análise de conteúdo, isto é, impõe-se uma obrigação às empresas de criarem setores especializados em averiguar se um conteúdo é ou não ilegal ou se fere os termos de serviços da plataforma, o que, por sua vez, significa a aplicação do Direito por empresas privadas - fato este que, como veremos, é uma das principais críticas dirigidas à nova lei alemã. Além disso, as plataformas podem ser instadas a agir por meio da criação de filtros de conteúdo e monitoramento dos discursos online, ou seja, *machine learning* capaz de detectar violações dos termos de serviço ou da lei, como no caso de direitos autorais.<sup>134</sup>

É importante destacar que, embora esse trabalho tenha um enfoque na questão da desinformação e da propagação de *fake news*, qualquer tipo de regulação imposta às plataformas tangencia outros “*harms*” praticados na esfera digital. Por esse motivo, eventuais regulações buscarão atingir esses

---

<sup>133</sup> Embora comentado no texto, nota-se que neste ponto reside a diferença entre os sistemas cuja responsabilização surge a partir do descumprimento da notificação judicial, entendida como a ideal para países que promovem a liberdade de expressão, ou a extrajudicial, criando casos em que as empresas poderiam vir a ser responsabilizadas após a solicitação de retirada do ar por um usuário qualquer, modelo este ligado a países com vestígios autoritários.

<sup>134</sup> Sobre o tema ver artigo de Clara Iglesias Keller disponível em <<https://www.jota.info/opiniao-e-analise/artigos/controle-de-conteudo-na-internet-filtros-de-upload-e-o-perigo-de-censura-previa-25062018>> (acessado em 04.02.19).

“males” em conjunto, pois todos envolvem restrições à liberdade de expressão em alguma medida (embora também estejam relacionados a questões de segurança cibernética e direitos de propriedade). Para direcionar o debate aos principais males a serem enfrentados em termos de discurso online, é necessário averiguar quais são as principais atividades dos usuários nas redes sociais que causam danos suficientemente grandes para que desejemos uma interferência legal, pelo menos em relação ao conteúdo e interação online, e como elas são atualmente reguladas.

Em última análise, para qualquer nação que esteja cogitando regular as redes sociais, caberá ao Executivo e ao Legislativo - como no caso brasileiro - decidir se existe necessidade de regulamentação adicional no domínio dos danos na Internet; e se isso requer, por exemplo, a criação de um novo órgão regulador e, se for esse o caso, sua forma de organização e alcance. Antes desses debates específicos, é importante sintetizar quais são os danos existentes, e quais atividades na Internet já estão (genericamente) regulamentadas e, portanto, mapear onde estariam as lacunas regulatórias. Este exercício pode ser feito ao se analisar a legislação de cada país.

Nota-se que os possíveis e prováveis danos possuem um espectro muito amplo, desde o conteúdo de menor potencial ofensivo que pode causar transtornos ou ofensas, como questões relacionadas ao *bullying*, indo até ao conteúdo extremista ou ilegal, afeto, por exemplo, aos discursos de ódio, racistas e xenófobos.<sup>135</sup> Neste ponto, é importante destacar que outros “temas sensíveis” já vêm ganhando maior atenção nos marcos regulatórios da esfera digital, como a divulgação de conteúdo pornográfico com o intuito de causar dano, ou melhor, o que se convencionou chamar de “pornografia de vingança”, e os direitos autorais e conexos. No caso brasileiro, a título de exemplo, o Marco Civil da Internet (Lei nº 12.965/14) trouxe obrigações aos provedores de aplicação de retirarem do ar conteúdo de pornografia de vingança ou que violem direitos autorais, sob pena de responsabilização. No caso do Marco

---

<sup>135</sup> Sobre o tema ver a série “Understanding Harmful Speech Online” do Berkman Klein Center de Harvard: FARIS, Robert. ASHAR, Amar. GASSER, Urs. JOO, Daisy, *Understanding Harmful Speech Online*. Berkman Klein Center Research Publication Nº. 2016-2. Disponível em: <<https://ssrn.com/abstract=2882824>> (acessado em 04.02.19).

Civil, tais exceções ao regime de responsabilidade subjetiva estão dispostas nos artigos 19, §2º, e 21.<sup>136</sup>

Sendo assim, as principais áreas de preocupação<sup>137</sup> visadas por vindouras regulações podem ser agrupadas em diferentes grupos, resumidos em: (i) conteúdo nocivo e ilegal, (ii) uso de dados privados; (iii) segurança cibernética e *hacking*, e (iv) questões relacionadas à interação segura com outras pessoas e comportamento humano.

Para lidar com essas questões a nível regulatório, em relação ao “uso de dados privados” e “segurança cibernética e *hacking*”, no caso brasileiro, a lacuna legal foi preenchida pela recente LGPD (Lei Geral de Proteção de Dados). Mas ainda precisaremos esperar para ver os efeitos de tal diploma legislativo no país, pois a norma está em período de vacância. No caso europeu, o GDPR (*General Data Protection Regulation*) também buscou tratar dos temas. É bom frisar que o marco legal europeu é anterior e serviu de inspiração para a confecção da LGPD brasileira, que copiou o modelo em vários aspectos.

Além desses, para encarar questões relacionadas a “conteúdo ilegal” normalmente os países vêm utilizando legislações existentes para espaços físicos, de maneira que a problemática está mais relacionada à execução (*enforcement*) dessas leis no mundo digital do que à criação de novas regras. Neste ponto, a lei alemã “NetzDG” procura justamente melhorar a aplicação da lei nas redes sociais visando enfrentar males próprios às mesmas.

É possível resumir esses danos em grupos:

- (A) ameaças prejudiciais (*harmful threats*) - declaração com intenção de causar dor, lesão, dano ou outra ação hostil, como intimidação;

---

<sup>136</sup> Art. 21 da Lei nº 12.965/14 - O provedor de aplicações de internet que disponibilize conteúdo gerado por terceiros será responsabilizado subsidiariamente pela violação da intimidade decorrente da divulgação, sem autorização de seus participantes, de imagens, de vídeos ou de outros materiais contendo cenas de nudez ou de atos sexuais de caráter privado quando, após o recebimento de notificação pelo participante ou seu representante legal, deixar de promover, de forma diligente, no âmbito e nos limites técnicos do seu serviço, a indisponibilização desse conteúdo.

<sup>137</sup> DENHAM, Elizabeth. *The future of Internet Harms Regulation*. Paper for the Digital, Culture, Media and Sport Committee. Supplementary written evidence submitted by Elizabeth Denham, Information Commissioner. Disponível em <http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/digital-culture-media-and-sport-committee/disinformation-and-fake-news/written/92853.html> (acessado em 04.02.19).

assédio psicológico, ameaças de natureza sexual, ameaças de morte, ameaças raciais ou religiosas conhecidas como crimes de ódio; hostilidade ou preconceito com base na raça, religião, orientação sexual, deficiência ou identidade transgênero de uma pessoa, ou seja, todas as formas de *hate speech*, podendo ser estendidas à misoginia;

- (B) danos econômicos (*economic harms*) - má conduta financeira; abuso de propriedade intelectual;
- (C) danos prejudiciais à segurança nacional (*harms to national security*) - extremismo violento, terrorismo, guerra cibernética patrocinada por Estados-nação estrangeiros;
- (D) danos emocionais (*emotional harms*) - prevenir danos emocionais sofridos por usuários de tal forma que não se acumulem até o limiar criminal de uma lesão psiquiátrica reconhecida. Por exemplo, através do abuso agregado de uma pessoa por muitos outros de uma forma que não aconteceria no mundo físico. Isso inclui, por exemplo, danos a pessoas vulneráveis - em relação a suicídio, anorexia, doença mental, etc.
- (E) danos à criança e ao adolescente (*harms to young people*) - intimidação, agressão, ódio, assédio sexual e comunicação, exposição a conteúdos nocivos ou perturbadores, higiene pessoal, abuso infantil (espécies de *bullyings* que caracterizariam crimes);
- (F) danos à Justiça e Democracia (*harms to justice and democracy*) - prevenir a intimidação de pessoas que participam do processo político, protegendo o processo cível e criminal.

No entanto, ainda existem vários danos na Internet que não parecem ser totalmente resolvidos por legislações e regulamentações já existentes, e por conter maiores lacunas - pelo menos no Brasil - exigem um pouco mais de atenção. Nesse sentido, quando se precisa lidar com a questão da disseminação de *fake news*, e da desinformação como um todo, entramos no espectro dos problemas relacionados a "conteúdo prejudicial" e "interação

segura com outras pessoas". Esses “*harms*” também são muito amplos e incluem, por exemplo:

↳ conduta online prejudicial (mas atualmente legal) de uma pessoa para outra, ou de muitas pessoas para uma pessoa. Isso pode incluir o abuso online do indivíduo (incluindo o *bullying* ou quando muitas pessoas "se juntam" contra um indivíduo e praticam abusos online) e também o comportamento inadequado (mas atualmente legal) para indivíduos vulneráveis, como crianças. As empresas de tecnologia geralmente tentam lidar com isso por meio de canais de denúncia e por meio da moderação automatizada de interações (inteligência artificial).

↳ publicação de conteúdos prejudiciais ou ofensivos (mas atualmente legais) feitos por usuários, especialmente onde pode ser consumido por indivíduos vulneráveis, como crianças. As empresas de tecnologia geralmente tentam lidar com isso por meio da moderação de conteúdo, usando uma mistura de filtragem de inteligência artificial e moderação de conteúdo humano, como a análise de conteúdo feita por empregados das próprias plataformas.

↳ problemas relacionados à exploração do comportamento humano, como design viciante, levando o usuário a passar excessivos períodos de tempo nas plataformas, ou outro tipo de criação de vícios comportamentais. As plataformas, muito impulsionadas pelo mercado competitivo e pela propagação de estudos associando o excessivo tempo nas redes sociais a questões de saúde mental, têm lidado com isso por meio de marketing pelo não uso dos *apps* (aplicativos) - movimento que foi chamado de "tempo bem gasto" -, havendo plataformas que disponibilizam relatórios, se requisitado, indicando o tempo gasto por dia nelas. De toda forma, isso não tem sido um grande foco



da autorregulação das empresas de tecnologia, pelo menos até recentemente.

Para suprir essas e outras lacunas regulatórias, que talvez ainda nem conheçamos ou que saibamos muito pouco a respeito, parece haver consenso de que qualquer abordagem coordenada aos danos na Internet deve incluir esforços governamentais e das próprias plataformas para melhorar a “alfabetização digital”, e isso deve incluir tornar o público mais consciente de seus direitos e proteções existentes, o que talvez não seja de muito interesse, economicamente falando, das plataformas. Certamente, o aprimoramento da educação digital contribuiria para lidar com alguns desses “*harms*” sem a necessidade de regulamentações ou reguladores adicionais.

No entanto, sabendo que a educação digital é uma solução em longo prazo, cumpre adentrar questões mais atinentes à regulação das plataformas como um todo. É preciso, então, recapitular o panorama regulatório dos provedores de aplicação, para argumentar que as plataformas não são irreguláveis, mas que as questões da Internet devem ser abordadas sempre que possível usando “técnicas do mundo físico”, e não obrigatoriamente pela criação de novos ilícitos cíveis e penais. Este norte tem sido amplamente considerado pelas políticas regulatórias nacionais e internacionais.

Nessa toada, alguns *insights* indicados por outros regimes regulatórios existentes, como, por exemplo, as comunicações, economia digital, saúde, segurança e meio ambiente, talvez possam fornecer algumas respostas. Isto porque, segundo pesquisadores da área, há semelhanças principiológicas que devem ser replicadas em qualquer que seja o regime regulatório pensado para as plataformas interativas.

Em primeiro lugar, aponta-se para o fato de que o mundo digital está em constante mudança, e que o Estado (ou entidade reguladora) não detém a expertise técnica, tampouco a informação que as empresas de tecnologia possuem, de maneira que quaisquer mudanças na política regulatória devem ocorrer de forma transparente e após consulta a uma série de partes interessadas; especialmente as plataformas. Neste ponto, é importante lembrar que uma das principais diferenças entre a regulação estatal e a autorregulação feita por entidades privadas reside justamente no fato de que,

em determinados segmentos do mercado, as empresas se encontram em uma posição técnica, tecnológica e de pesquisa muito mais avançada do que organismos estatais, de sorte que as decisões tendem a estar mais bem embasadas cientificamente.<sup>138</sup> Em segundo lugar, é crucial que os players envolvidos, ou seja, as plataformas e as agências reguladoras, tenham alguma forma de supervisão e fiscalização - incluindo penalidades criminais. Em terceiro lugar está algo que vem sendo amplamente debatido no Brasil: também é muito importante que os reguladores sejam independentes do governo (maiorias ocasionais) e da indústria (mercado).<sup>139</sup>

Em suma, vislumbra-se um modelo regulatório em que a violação do dever legal pode levar à responsabilização das plataformas, mas no qual as questões de padrões a serem seguidos não são deixadas apenas nas mãos da indústria ou dos governos, de modo que haja transparência e que os procedimentos internos sigam regras de ordem pública, respeitando o contraditório das partes, diminuindo as chances de *overblocking*, e não penalizando as plataformas por questões impossíveis de serem resolvidas devido à arquitetura do código.

No entanto, para lidar com essas e outras dificuldades, é possível extrair alguns princípios importantes de outros regimes regulatórios, que sintetizam o norte das propostas regulatórias, como o princípio da precaução (*precautionary principle*) e o dever de cuidar (*duty to care*).

---

<sup>138</sup> BINENBOJM, Gustavo. *Poder de polícia, ordenação, regulação: transformações político-jurídicas, econômicas e institucionais do direito administrativo ordenador*, op. cit. nota 105.

<sup>139</sup> A Autoridade Nacional de Proteção de Dados (ANPD), entidade prevista no vetado artigo 55 da Lei nº 13.709/2018 (LGPD) foi criada por meio de medida provisória (MP 869) em 20 de dezembro de 2018, nos últimos dias do governo Temer. A autoridade será a responsável pela fiscalização do tratamento de dados pessoais no país, podendo aplicar sanções, no entanto havia a expectativa de que a ANPD fosse criada via decreto que estabelecesse a autoridade como uma autarquia, com independência financeira, orçamento próprio e realização de sabatinas para a seleção de diretores e membros de conselho. Porém, com a MP de Michel Temer, a ANPD fica estabelecida como uma autoridade vinculada à Presidência da República, sem orçamento independente, e com reduzida autonomia. O veto à criação da autoridade pelo Legislativo se deu em razão de inconstitucionalidade por vício de iniciativa, pois a instituição de uma autarquia por aquele Poder geraria gastos ao Executivo.

### **3.2 O princípio da precaução e o dever de cuidado como nortes regulatórios: é possível exigir um dever de monitoramento das plataformas?**

Como exposto anteriormente, a problemática posta é que as redes sociais são diferentes de muitos outros negócios em setores comparáveis, como telecomunicações ou radiodifusão, e, com isso, as possibilidades regulatórias também são severamente alteradas. Portanto, também há uma fundamental diferença entre regulações do “mundo físico” e regulações do “mundo digital” no que tange à sua forma de implementação. Como Lawrence Lessig argumentou, o código de computador define as condições nas quais a Internet (e todos os computadores) funciona. O código é a arquitetura do ciberespaço e afeta o que as pessoas fazem online: o código permite, facilita e às vezes proíbe. Portanto, a primeira coisa que devemos ter em mente é o fato de que eventuais regulações precisam ser possíveis levando em conta a estrutura do código e as características arquitetônicas da Internet. Nas palavras de Lessig:

O código básico da Internet implementa um conjunto de protocolos chamado TCP / IP. Esses protocolos permitem a troca de dados entre redes interconectadas. Essa troca ocorre sem que as redes saibam o conteúdo dos dados, ou sem qualquer idéia verdadeira de quem na vida real é o remetente de um determinado dado. Esse código é neutro acerca dos dados e ignorante sobre o usuário. Esses recursos do TCP/IP têm consequências para a "regulabilidade" do comportamento na Internet. Eles dificultam o comportamento regulador. Na medida em que é difícil identificar quem são as pessoas, é mais difícil traçar o comportamento de um indivíduo em particular. E, na medida em que é difícil identificar que tipo de dado está sendo enviado, é mais difícil regular o uso de determinados tipos de dados. Essas características arquitetônicas da Internet significam que os governos são relativamente deficientes em sua capacidade de regular o comportamento na Internet.<sup>140</sup>

---

<sup>140</sup> Tradução livre do autor do trecho “*The basic code of the Internet implements a set of protocols called TCP/IP. These protocols enable the exchange of data among interconnected networks. This exchange occurs without the networks knowing the content of the data, or without any true idea of who in real life the sender of a given bit of data is. This code is neutral about the data, and ignorant about the user. These features of TCP/IP have consequences for the "regulability" of behavior on the Internet. They make regulating behavior difficult. To the extent that it is hard to identify who people are, it is harder to trace behavior back to a particular individual. And to the extent it is hard to identify what kind of data is being sent, it is harder to*

Não há dúvida de que o trabalho de Lawrence Lessig está para a regulação de tecnologias emergentes como o trabalho de Habermas está para os estudos sobre esfera pública, funcionando como uma espécie de trabalho seminal, referência para qualquer reflexão sobre o assunto. No entanto, o precioso caminho aberto por Lessig talvez não satisfaça plenamente as necessidades de atores legais que estão tendo que lidar com problemas reais, e começando a elaborar eventuais regulações para esses problemas emergentes do mundo digital. Assim, a extensão e gravidade dos danos causados requerem medidas rápidas e eficazes. Em analogia à expressão cunhada pelo catedrático, segundo o qual “*code is law*”, o que eventuais regulações à Internet tentarão fazer é, na verdade, transformar “*law in code*”, a lei em código, o que não é possível em todos os casos, devido à própria arquitetura da Internet.

Para Lawrence Lessig, a regulação de novas tecnologias é uma tarefa difícil que exige uma diversidade de setores (com “modalidades regulatórias” distintas). O professor coloca quatro delas: (i) a lei, (ii) o mercado, (iii) as normas sociais e (iv) a própria tecnologia (também cunhada como a arquitetura ou o “código”). Assim, não haveria uma regulação em geral, mas apenas modos específicos de regulação: pela esfera jurídica, econômica, social e tecnológica/arquitetônica. Portanto, regular novas tecnologias significa construir algum tipo de concordância ou interação optimal entre essas quatro diferentes modalidades de regulação. Lessig coloca que o que deve ser buscado é uma boa forma de concordância - isto é, uma forma de concordância que satisfaça os objetivos apropriados das diferentes práticas e setores envolvidos. Essa boa concordância, ou interação das modalidades, é o que ele chama de “*optimal mix*”.<sup>141</sup>

---

*regulate the use of particular kinds of data. These architectural features of the Internet mean that governments are relatively disabled in their ability to regulate behavior on the Net.” in LESSIG, Lawrence. Code is Law. On Liberty in Cyberspace. Harvard Magazine. Disponível em <<http://harvardmagazine.com/2000/01/code-is-law-html>>. Trata-se de resenha da famosa obra LESSIG, Lawrence. Code. Version 2.0. New York: Basic Books. 2006.*

<sup>141</sup> GUTWIRTH, Serge. DE HERT, Paul. DE SUTTER, Laurent. *The trouble with technology regulation from a legal perspective. Why Lessig’s ‘optimal mix’ will not work.* Oxford Regulating Technologies. Legal Futures, Regulatory Frames and Technological Fixes (2008). Disponível em <[http://works.bepress.com/serge\\_gutwirth/1/](http://works.bepress.com/serge_gutwirth/1/)> (acessado em 04.02.19).

No Direito Administrativo brasileiro, Patrícia Baptista e Clara Iglesias Keller bem resumiram as lições de Lessig no artigo “*Por que, quando e até onde regular as novas tecnologias? Entre inovação e preservação, os desafios trazidos pelas inovações disruptivas*”:

Aplicando cada uma dessas modalidades às peculiaridades da Internet, o autor conclui que, nesse ambiente, a modalidade de regulação por *arquitetura* se destaca, enquanto as outras formas de determinar comportamentos se revelam menos eficazes. A arquitetura dos sistemas - determinado pelos códigos que os constroem - seria a forma mais eficaz de estimular ou prevenir comportamentos. Destaque-se que, para o autor, cada modalidade de regulação impõe um tipo de custo diferente sobre os agentes regulados, com o objetivo de que se eles se engajem no comportamento desejado, de forma que as quatro modalidades interagem de forma dinâmica e simultânea (sendo o código, no caso da Internet, a modalidade mais promissora).<sup>142</sup>

No caso da regulação das plataformas, de um lado temos a dificuldade em regular tecnologias emergentes, que exigem ações multidisciplinares, e por isso devem estar embasadas em constatações científicas e técnicas de inúmeras áreas do conhecimento, o que demanda tempo. Por outro lado, temos urgência, pois os danos causados pelas e nas redes sociais são cada vez mais aparentes, já sendo certo que o comportamento humano está sendo modificado como resultado das mudanças introduzidas pelas vidas online e pelas Sociedades Algorítmicas. As *fake news* - como gênero -, da espécie desinformação, são apenas alguns *harms* da atualidade, os quais colocam em choque a democracia e a liberdade de expressão.

A conclusão, portanto, é que a percepção de graves danos na sociedade encurta nossa possibilidade de espera por respostas espontâneas e naturais da comunidade, indicando que soluções regulatórias baseadas em risco devem ser adotadas. Agora, se somarmos a isto o fato de que a modalidade arquitetônica se mostra a tipologia regulatória mais eficiente dentro da teoria regulatória das tecnologias, produzindo melhores resultados, há motivos para crer que a implementação de regulações por código também serão buscadas. Como se sabe, tais operações devem ser implementadas pelas próprias

---

<sup>142</sup> BAPTISTA, Patrícia; KELLER, Clara Iglesias. *Por que, quando e como regular as novas tecnologias? Os desafios trazidos pelas inovações disruptivas*. Revista de Direito Administrativo, Rio de Janeiro, v. 273, p. 123-163, set. 2016. ISSN 2238-5177. Disponível em: <<http://bibliotecadigital.fgv.br/ojs/index.php/rda/article/view/66659>> (acessado em 04.02.19).

plataformas, de maneira que acabam sendo nada mais do que as plataformas se autogovernando, o que poderia acontecer por alterações nos termos de serviço, ou pela criação de filtros e inteligência artificial capazes de detectar a ocorrência de danos. Percebemos, assim, que a autorregulação e a autogovernança das plataformas já acontecem. A pergunta que fica é: como acontecem?

É nesse contexto de urgência social, e de incertezas científicas, que entram em cena mecanismos regulatórios que se propõem a antecipar e limitar as consequências prejudiciais à sociedade antes que elas ocorram - são mecanismos baseados no risco da atividade. Por essa razão, pesquisadores vêm indicando que a melhor forma de lidar com essas dificuldades trazidas pelas tecnologias disruptivas é procurar respostas dentro das legislações já existentes, adequando a teoria - já posta em prática pelo Estado e pelo mercado em outros setores - às possibilidades arquitetônicas da Internet. Assim, qualquer resposta regulatória para reduzir os danos online - a desinformação inclusa -, precisa, em primeiro lugar, respeitar as possibilidades técnicas do setor: ou seja, compreender o que é, de fato, factível e factível pelos agentes regulados; em segundo lugar, deve buscar estar embasada o máximo possível em leis já existentes, para ter que lidar o mínimo possível com a criação de novas espécies legais, que talvez não se adequem ao modelo arquitetônico (código) e para que seja uma intervenção minimalista no setor que ainda está em desenvolvimento, evitando impedir a inovação.

Ora, a formulação de políticas regulatórias tradicionais exige que as decisões políticas sejam informadas por evidências objetivas rigorosamente estabelecidas. Normalmente, a ação em uma questão só é tomada após a consulta e coleta de tais evidências. Mas, em áreas inovadoras, muitas vezes não há pesquisa científica de longo prazo; ou tal evidência chega tarde demais para fornecer uma medida eficaz contra os danos já percebidos e os prováveis. A propagação rápida de serviços de mídia social, sujeita a ondas de moda, é um desafio particular para a necessidade de se obter evidências objetivas em longo prazo, pois os serviços são recentes e estão em constante mutação. Diante de tal incerteza científica, há mecanismos regulatórios desenvolvidos em outros setores justamente para lidar com cenários de incerteza científica e urgência social, a exemplo do “princípio da precaução” e do “dever de cuidado”.

O princípio da precaução fornece uma base para a prevenção de danos baseados em riscos. Para modelos regulatórios de áreas como meio ambiente ou saúde, o princípio da precaução deve ser aplicado quando, com base nas melhores opiniões científicas disponíveis no prazo para a tomada de decisões, existirem boas razões para crer que podem ocorrer efeitos prejudiciais para a saúde humana, animal ou vegetal, ou para o meio ambiente. Portanto, percebe-se que o princípio vem à tona quando evidências precoces de dano ao público são aparentes, mas antes que pesquisas científicas inequívocas tenham constatado o prejuízo em si.

Em suma, para prevenir novos danos, abre-se mão da necessidade da certeza científica, bastando que o risco seja suficientemente alto para regular. Remonta-se o conteúdo jurídico do princípio da precaução à “Carta Mundial da Natureza”, de 1982, cujo princípio nº 11, “b”, estabelece a necessidade de os Estados controlarem as atividades potencialmente danosas ao meio ambiente, ainda que seus efeitos não sejam completamente conhecidos. Posteriormente, no Direito brasileiro, o princípio da precaução foi introduzido como Princípio XV pela Declaração do Rio de Janeiro sobre Meio Ambiente e Desenvolvimento (Eco-92).<sup>143</sup> Além desses documentos, o princípio da precaução estaria contido no art. 225, § 1º, IV e V da Constituição.<sup>144</sup> Também decorre do art. 54<sup>145</sup> da Lei nº 9.605/98, que dispõe sobre sanções derivadas de atividades lesivas ao meio ambiente.

---

<sup>143</sup> Declaração do Rio sobre Meio Ambiente e Desenvolvimento - Princípio 15 - Com o fim de proteger o meio ambiente, o princípio da precaução deverá ser amplamente observado pelos Estados, de acordo com suas capacidades. Quando houver ameaça de danos graves ou irreversíveis, a ausência de certeza científica absoluta não será utilizada como razão para o adiamento de medidas economicamente viáveis para prevenir a degradação ambiental. Disponível em: <[http://www.meioambiente.pr.gov.br/arquivos/File/agenda21/Declaracao\\_Rio\\_Meio\\_Ambiente\\_Developolvimento.pdf](http://www.meioambiente.pr.gov.br/arquivos/File/agenda21/Declaracao_Rio_Meio_Ambiente_Developolvimento.pdf)> (acessado em 04.02.19).

<sup>144</sup> Constituição Federal. Art. 225. § 1º - Para assegurar a efetividade desse direito, incumbe ao Poder Público: ... IV - exigir, na forma da lei, para instalação de obra ou atividade potencialmente causadora de significativa degradação do meio ambiente, estudo prévio de impacto ambiental, a que se dará publicidade; V - controlar a produção, a comercialização e o emprego de técnicas, métodos e substâncias que comportem risco para a vida, a qualidade de vida e o meio ambiente

<sup>145</sup> Art. 54 da Lei n. 9605/98 - Causar poluição de qualquer natureza em níveis tais que resultem ou possam resultar em danos à saúde humana, ou que provoquem a mortandade de animais ou a destruição significativa da flora: § 3o Incorre nas mesmas penas previstas no parágrafo anterior quem deixar de adotar, quando assim o exigir a autoridade competente, medidas de precaução em caso de risco de dano ambiental grave ou irreversível.

Na Repercussão Geral do RE 627189/SP, o Supremo Tribunal Federal entendeu que decorre da inteligência do princípio da precaução

a constatação de que a evolução científica poderia trazer riscos, muitas vezes imprevisíveis ou imensuráveis, a exigir uma reformulação das práticas e procedimentos tradicionalmente adotados na respectiva área da ciência. Apontou que o princípio da precaução não prescindiria de outros elementos considerados essenciais para uma adequada decisão estatal, a serem observados sempre que estiver envolvida a gestão de riscos: a) a proporcionalidade entre as medidas adotadas e o nível de proteção escolhido; b) a não discriminação na aplicação das medidas; e, c) a coerência das medidas que se pretende tomar com as já adotadas em situações similares ou que utilizem abordagens similares. Portanto, na aplicação do princípio da precaução a existência de riscos decorrentes de incertezas científicas não deveria produzir uma paralisia estatal ou da sociedade. Por outro lado, a aplicação do princípio não poderia gerar como resultados temores infundados. Assim, em face de relevantes elementos de convicção sobre os riscos, o Estado deveria agir de forma proporcional. Por sua vez, o eventual controle pelo Poder Judiciário quanto à legalidade e à legitimidade na aplicação desse princípio haveria de ser realizado com prudência, com um controle mínimo, diante das incertezas que reinam no campo científico.<sup>146</sup>

Percebe-se, então, que, de um lado, o princípio da precaução autoriza a ação estatal em setores potencialmente prejudiciais e com incertezas científicas, mas, por outro lado, ele também serve de norte para balizar as medidas e sanções adotadas dentro de parâmetros proporcionais a essas incertezas e riscos.

É diante desse cenário que outro mecanismo regulatório alternativo surge para contribuir com a averiguação da cautela adotada pelos agentes regulados na contenção dos riscos e danos, o princípio do *duty to care*, que pode ser traduzido como “dever de diligência”, “dever de cuidado” ou “dever de

---

<sup>146</sup> “Essa a tese que, por maioria, o Plenário fixou para efeito de repercussão geral ao dar provimento, por maioria, a recurso extraordinário para julgar improcedentes pedidos formulados em ações civis públicas. Essas ações debateram o direito fundamental à distribuição de energia elétrica, ao mercado consumidor, de um lado, e o direito à saúde daqueles que residem em locais próximos às linhas pelas quais se efetua a transmissão, de outro. Na espécie, acórdão de tribunal de justiça estadual impusera obrigação de fazer a concessionária de serviço público no sentido de observar padrão internacional de segurança e, em consequência, reduzir campo eletromagnético em suas linhas de transmissão de energia elétrica. A decisão recorrida fundamentara-se no princípio da precaução e no direito fundamental a um meio ambiente ecologicamente equilibrado e à sadia qualidade de vida. O Plenário reafirmou que a proteção do meio ambiente e da saúde pública com desenvolvimento sustentável seria obrigação constitucional comum a todos os entes da Federação.”  
Repercussão Geral RE 627189/SP, rel. Min. Dias Toffoli, 8.6.2016. (RE-627189). Informativo 829 do STF, disponível em <http://www.stf.jus.br/arquivo/informativo/documento/informativo829.htm#repercussao> > (acessado em 04.02.19).



cuidar”, possuindo nomenclaturas diferentes a depender do seu uso em cada ramo do direito.

O “dever de cuidado” é simples, amplo e condicionado à “prova científica futura”.<sup>147</sup> Está mais focado no objetivo buscado do que nos detalhes dos meios utilizados por aqueles que estão em melhor posição para encontrar soluções adequadas ao contexto - aqueles que estão sujeitos ao dever de cuidar -, no caso, os agentes regulados (as plataformas). O *duty to care* é mais direto: uma pessoa (ou uma empresa) com “dever de diligência” deve tomar cuidados necessários em relação a uma atividade particular, uma vez que a atividade afeta (negativamente) pessoas ou ambientes específicos. Se esse “ente responsável” não tomar os devidos cuidados e alguém sofrer danos, pode haver consequências legais, pois a empresa lida com “atividade de risco” (podendo os riscos ser variáveis) e, por isso, deve ser suficientemente diligente para evitá-los. Por outro lado, deve-se ter em mente que um hipotético dever de cuidado não requer perfeição na atuação, tendo muito mais a ver com a averiguação de se os devidos cuidados foram suficientemente tomados pelo agente regulado para minimizar a ocorrência de danos, considerando os riscos existentes e segundo análises científicas e epistêmicas não tão conclusivas.

Outra característica relevante do *duty to care* é que o elemento preventivo reduz o sofrimento das potenciais vítimas, podendo também impedir que os comportamentos atinjam um limiar criminal, porque, caso tenham sido adotadas todas as medidas possíveis para a contenção de um dano, no caso de sua ocorrência o ônus probatório da culpa (ou dolo) aumenta - mantendo-se, por outro lado, a responsabilidade civil objetiva por risco da atividade.

A transposição da lógica do dever de cuidar para o mundo digital também tem um aspecto interessante - que será melhor abordado quando

---

<sup>147</sup> Nesse sentido: “*Duties of care are expressed in terms of what they want to achieve – a desired outcome (ie the prevention of harm) rather than necessarily regulating the steps – the process – of how to get there. This fact means that duties of care work in circumstances where so many different things happen that you couldn’t write rules for each one. This generality works well in multifunctional places like houses, parks, grounds, pubs, clubs, cafes, offices and has the added benefit of being to a large extent futureproof.*” in “Reducing harm in social media through a duty of care”, um dos artigos dentro da série de artigos sobre as propostas regulatórias levadas ao Parlamento do Reino Unido sobre regulação das redes sociais por William Perrin e Lorna Woods, disponível em <<https://www.carnegieuktrust.org.uk/blog/reducing-harm-social-media-duty-care/>>. A série está disponível em <<https://www.carnegieuktrust.org.uk/project/harm-reduction-in-social-media/?blog=view>> (acessados em 04.02.19).

falamos da evolução jurisprudencial da *public forum doctrine*, ou seja, das restrições à liberdade de expressão em lugares públicos - ligado à ideia de que, para o trato regulatório, as redes sociais mais parecem lugares públicos no sentido de fóruns públicos, do que a imprensa.

Resumindo, uma comparação comumente feita por pesquisadores é que os serviços de mídia social seriam como editores, porquanto o tratamento algorítmico que fazem para disponibilizar conteúdo aos usuários se assemelha ao controle editorial feito pela imprensa, setor no qual regulamentações próprias foram construídas ao longo do século XX. Esses temas foram abordados na primeira parte do trabalho, especialmente quando explicamos o sentido de Sociedades Algorítmicas, personalização do conteúdo e filtros-bolha. Acontece que, para outros pesquisadores, especialmente os que focam em questões relacionadas ao comportamento humano nas redes, no que tange à redução de danos praticados na esfera digital, a principal analogia a ser utilizada para as redes sociais é a de que elas devem ser vistas como um local público, ou melhor, um reino com direitos de propriedade privada, mas com características de lugar público - como um escritório, um bar ou um parque temático. Isto porque bilhões de pessoas vão às redes sociais - propriedades de empresas de tecnologia - para fazer uma vasta gama de coisas diferentes que, por sua vez, afetam outros usuários e a sociedade de maneiras distintas. Portanto, sob o enfoque relacional, e não mais sob a ótica do controle de conteúdo, as plataformas interativas mais parecem propriedades privadas com características de lugares públicos do que editoriais dentro da imprensa.

A lei provou ser muito boa para regular esses tipos de proteção em espaços físicos de propriedade. Por exemplo: espaços de trabalho ou espaços públicos pertencentes ou fornecidos por empresas devem ser locais seguros para as pessoas que os utilizam, decorrendo daí a inteligência do *duty to care*. A lei impõe um "dever de cuidado" aos proprietários desses espaços, de sorte que eles devem tomar medidas razoáveis para evitar danos, como, a exemplo, regras de segurança do trabalho. Embora a empresa tenha liberdade para adotar sua própria abordagem para cada questão, a noção do que é "razoável" está sujeita à supervisão de um regulador, ou à combinação de órgãos reguladores e fiscalizadores, mantendo-se sempre a possibilidade de acesso ao Judiciário em caso de litígio.

Dentro desses sistemas, se o dano acontecer e restar comprovada a falta do devido cuidado, à vítima poderá ser reconhecido o direito à reparação, além de qualquer sanção que um regulador possa tomar. Ao “obrigar” as empresas a investirem em segurança, o mercado funciona melhor, pois o agente regulado arca com os custos totais de suas ações, em vez de obter um “subsídio implícito”, enquanto a sociedade arca com os custos da atividade. Pensemos, por exemplo, no caso dos custos previdenciários impostos à sociedade por causa de acidentes de trabalho originados de condutas não diligentes do empregador.

Como indica Ana Frazão na série de artigos<sup>148</sup> sobre a Lei Geral de Proteção de Dados, tal diploma, além de prever o princípio da prevenção (art. 6, VIII), também fixa o dever de cuidado, a exemplo do art. 14, § 5º (“todos os esforços razoáveis”). Acredita-se ser possível extrair também um dever de cuidado, ou dever de diligência, da inteligência dos artigos 46,<sup>149</sup> 49,<sup>150</sup> e 50<sup>151</sup> e seus parágrafos.

No caso dos danos da Internet e da governança, as plataformas poderiam ter abordagens diferentes para cada tipo de “*harm*”, umas mais brandas e outras mais severas, e, com isso, cria-se também a possibilidade de surgirem mercados para grupos específicos, com diferentes níveis e tipos de salvaguardas. Ou seja, vislumbra-se que eventuais padrões de conteúdo, fala

---

<sup>148</sup> Primeira e sexta partes da série sobre a LGPD, disponível em: <<https://www.jota.info/autor/ana-frazao>> (acessado em 04.02.19).

<sup>149</sup> Art. 46. Os agentes de tratamento devem adotar medidas de segurança, técnicas e administrativas aptas a proteger os dados pessoais de acessos não autorizados e de situações acidentais ou ilícitas de destruição, perda, alteração, comunicação ou qualquer forma de tratamento inadequado ou ilícito.

<sup>150</sup> Art. 49. Os sistemas utilizados para o tratamento de dados pessoais devem ser estruturados de forma a atender aos requisitos de segurança, aos padrões de boas práticas e de governança e aos princípios gerais previstos nesta Lei e às demais normas regulamentares.

<sup>151</sup> Art. 50. Os controladores e operadores, no âmbito de suas competências, pelo tratamento de dados pessoais, individualmente ou por meio de associações, poderão formular regras de boas práticas e de governança que estabeleçam as condições de organização, o regime de funcionamento, os procedimentos, incluindo reclamações e petições de titulares, as normas de segurança, os padrões técnicos, as obrigações específicas para os diversos envolvidos no tratamento, as ações educativas, os mecanismos internos de supervisão e de mitigação de riscos e outros aspectos relacionados ao tratamento de dados pessoais. § 1º Ao estabelecer regras de boas práticas, o controlador e o operador levarão em consideração, em relação ao tratamento e aos dados, a natureza, o escopo, a finalidade e a probabilidade e a gravidade dos riscos e dos benefícios decorrentes de tratamento de dados do titular. § 2º Na aplicação dos princípios indicados nos incisos VII e VIII do caput do art. 6º desta Lei, o controlador, observados a estrutura, a escala e o volume de suas operações, bem como a sensibilidade dos dados tratados e a probabilidade e a gravidade dos danos para os titulares dos dados, poderá:

ou comportamento que não são aceitáveis em uma determinada plataforma, podem encontrar guarida em outras plataformas. Os usuários sempre podem, é claro, participar de mais de um serviço, se assim desejarem.

Uma importante observação a ser feita é que autores no tema têm sugerido a aplicação desse modelo apenas para as maiores plataformas porque, assim, mantém-se a possibilidade de surgimento de novas empresas, preservando-se também a liberdade de expressão o máximo possível. Como explicado, prevê-se que as plataformas possam ter abordagens diferentes, até mesmo porque suas arquiteturas são diferentes,<sup>152</sup> e que um mercado possa surgir em plataformas que visem a grupos específicos, que desejam muito praticar esse ou aquele tipo de conduta (legal), porém tida como prejudicial por determinado termo de serviço. Em suma, padrões de conteúdo ou fala que não são aceitáveis em uma plataforma podem encontrar hospedagem em outro lugar, maximizando a liberdade de expressão e mantendo outros canais abertos.<sup>153</sup> Além disso, aplicar regulações mais severas apenas às maiores plataformas resguarda a inovação no mercado, não sobrecarregando plataformas com menos recursos.<sup>154</sup>

Aqui, é de extrema relevância levantar um debate que, embora merecesse ser aprofundado, infelizmente não será possível devido ao escopo do trabalho. Trata-se do debate sobre o mercado das redes sociais se aproximarem muito de oligopólios,<sup>155</sup> especialmente o mercado publicitário da Internet (anúncios). Nesse sentido, tratando-se de um oligopólio ou não, o fato é que as grandes empresas de tecnologia têm muito poder e operam em um

---

<sup>152</sup> Por exemplo, das grandes plataformas, o Snapchat parece não ser uma plataforma que sofra com grandes problemas de disseminação de *fake news*. Sobre o tema ver <<https://www.bloomberg.com/news/features/2017-10-26/how-snapchat-has-kept-itself-free-of-fake-news>> (acessado em 04.02.19).

<sup>153</sup> Essa solução provavelmente criará nichos, também explorados economicamente, com seus próprios termos de uso e regras de comunidade. Por exemplo, existem redes sociais para pessoas interessadas em certos tipos de fetiches sexuais, com verificações específicas para controle de idade, exposição, etc.

<sup>154</sup> Nesse sentido, por exemplo, a recente NetzDG alemã para a melhoria na aplicação da lei nas redes sociais libera de determinadas obrigações às plataformas interativas que possuem, domesticamente, menos de 2 milhões de usuários registrados [artigo 1º, §1º (2)].

<sup>155</sup> E nem poderia ser um oligopólio de fato e operar no Brasil, pois, não obstante o tema resgatar o debate sobre as plataformas interativas serem ou não “meios de comunicação” em sentido formal da lei, segundo o artigo 220, § 5º da Constituição, os meios de comunicação social não podem, direta ou indiretamente, ser objeto de monopólio ou oligopólio: CRFB/88 - art. 220, §5º - § 5º Os meios de comunicação social não podem, direta ou indiretamente, ser objeto de monopólio ou oligopólio.

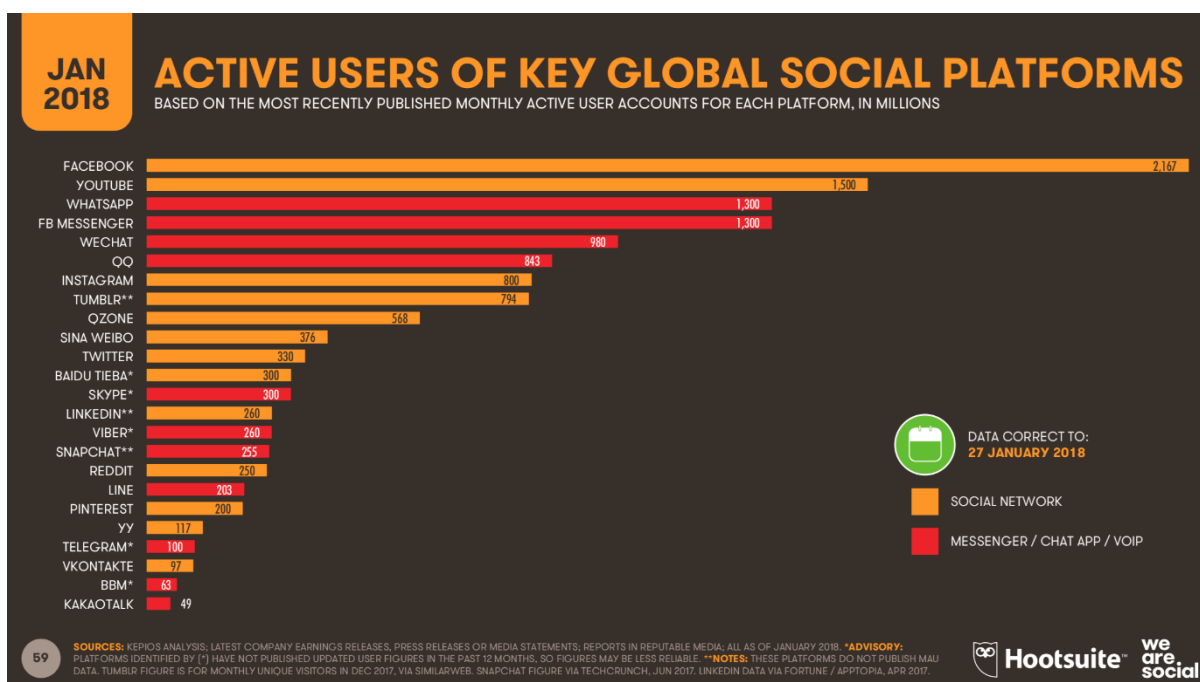
ambiente altamente desregulado, e talvez argumentos embasados em visões liberais sobre competição seriam uma ilusão.<sup>156</sup> O Facebook, por exemplo, é o maior controlador de plataformas interativas, como o próprio Facebook, mas também o Instagram e o WhatsApp. Vejamos alguns dados da We Are Social, uma agência de marketing digital,<sup>157</sup> que rapidamente ilustram o debate: das cerca de 4 bilhões de pessoas com acesso à Internet, aproximadamente 3.2 bilhões são usuárias ativas de redes sociais.



Mesmo considerando que muitos usuários possuem mais de uma plataforma, nota-se também que o mercado é controlado por poucas Intermediárias da Internet, embora elas prefiram ser chamadas de “empresas de tecnologia”:

<sup>156</sup> Sobre o tema ver: <<http://www.pewinternet.org/2018/04/17/the-future-of-well-being-in-a-tech-saturated-world/>>; <<http://www.pewinternet.org/2017/10/19/the-future-of-truth-and-misinformation-online/>> (acessados em 04.02.19).

<sup>157</sup> The global state of digital in 2018 - from Argentina to Zambia. Slides 7 e 59. Disponível em: <<https://hootsuite.com/pt/pages/digital-in-2018>> (acessado em 04.02.19).



Sendo assim, feitas todas as considerações sobre a dificuldade de regular o setor, bem como sobre a necessidade de fazê-lo mantendo-se a possibilidade de inovação, a generalidade e a simplicidade do *duty to care* funcionam bem em confronto à amplitude, complexidade e rápido desenvolvimento das redes sociais, onde escrever “regras de convivência” detalhadas em lei é tão impossível quanto coletar evidências com rapidez suficiente para fornecer uma base científica suficientemente completa para sustentar intervenções regulatórias severas. Por esse motivo, tornar proprietários e operadores dos maiores serviços de mídia social responsáveis pelos custos e ações de redução de danos - e assim internalizar esses custos para o provedor (em vez de gerar custos externos para a sociedade) - também poderá ajudar o mercado a funcionar melhor.

Acontece que, no caso dos danos online, há outros dois debates que devem ser considerados. O primeiro é que, exigir um dever de cuidado das plataformas para que governem seus “reinos digitais” da maneira mais diligente possível pode significar a exigência de certo “dever de monitoramento” dos conteúdos disponibilizados por terceiros, o que se mostra uma dificuldade tanto pelo ponto de vista da viabilidade técnica de fazê-lo, quanto por questões constitucionais ligadas à censura. Isto porque uma das formas pela qual as plataformas reconhecem discursos proibidos pelos seus termos de uso se dá

pelo uso de *machine learning* e inteligência artificial, além dos procedimentos de denúncia por outros usuários. Tendo isto em vista, é importante lembrar os ensinamentos de Lessig, segundo os quais eventuais regulações impostas ao mundo digital devem respeitar a arquitetura do código e serem tecnicamente viáveis, o que, sob um aspecto, também significa a viabilidade econômica de serem criados procedimentos empresariais internos de averiguação dos discursos. Sendo assim, exigir um dever de monitoramento esbarra em dificuldades técnicas, econômicas e democráticas, posto que, do monitoramento à censura, a fronteira é breve e o terreno é perigoso.

E, na verdade, nem precisaria ser, porque a proibição da censura costuma estar assegurada constitucionalmente de forma bastante privilegiada nas democracias modernas, como no caso brasileiro.

A Constituição Federal expressamente prevê a vedação da censura nos artigos 5º, IX <sup>158</sup> e 220, §2º. <sup>159</sup> Ao tocarmos no tema em questão, logo nos remetemos aos tempos de regime militar, quando se tratava de prática exercida por órgãos administrativos governamentais. No entanto, essas técnicas arcaicas de censura direta por governos deram lugar a formas muito mais sutis e tecnológicas de controle; agora, tanto por governos, quanto por atores privados. Neste ponto, no artigo “Atuação e responsabilidade dos provedores diante das *fake news* e da desinformação”, André Zonaro Giacchetta explica que

a vedação da censura (...) incide sobre diversos tipos de manifestações, inclusive a informação jornalística, e não se limita a motivações políticas, ideológicas e artísticas. Enquadram-se à vedação prevista nesse dispositivo os atos de controle e proibição da veiculação da informação que não advenham da realização de uma ponderação entre a liberdade de informação e direitos fundamentais circunstancialmente contrapostos, revelando-se, por conseguinte, arbitrários, imponderados, dissociados dos legítimos fins da Constituição Federal. A vedação constitucional à censura da liberdade de informação (inclusive jornalística) não se limita às hipóteses de exclusão de determinado conteúdo da Internet.

---

<sup>158</sup> CRFB/88 - 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: (...) IX - é livre a expressão da atividade intelectual, artística, científica e de comunicação, independentemente de censura ou licença;

<sup>159</sup> CRFB/88 - Art. 220. A manifestação do pensamento, a criação, a expressão e a informação, sob qualquer forma, processo ou veículo não sofrerão qualquer restrição, observado o disposto nesta Constituição. (...) § 2º É vedada toda e qualquer censura de natureza política, ideológica e artística.

Pretender dificultar o acesso do público a tal conteúdo, por exemplo, obstando sua localização a partir de uma ferramenta de busca na Web, nada mais é do que cercear, de modo conveniente aos interesses do particular, o direito constitucional de informação.<sup>160</sup>

Assim, percebe-se que a vedação à censura possui facetas distintas, podendo significar “direito de informar”, que assegura ao indivíduo a faculdade de se comunicar, mas também o “direito de se informar”, segundo o qual todos têm direito de acesso à informação. Observa-se que o direito de se informar pode ser caracterizado como direito coletivo, ou seja, percebem-se os titulares do direito não apenas como pessoas isoladamente consideradas, mas a coletividade como um todo.

Portanto, considerando que a moderação de conteúdo pelas plataformas interativas pode acontecer após a publicação do conteúdo (por denúncia de outros usuários) ou na modalidade preventiva, onde há censura de conteúdo antes da publicação (por filtros), sem dúvida, a segunda se mostra muito mais desafiadora sob o ponto de vista técnico, pois as plataformas ficariam encarregadas de moderar um volume enorme, em um fluxo constante de conteúdo gerado por usuários de todo o mundo.

Mesmo a tarefa podendo ser automatizada até certo ponto, o julgamento humano ainda é necessário por duas razões. Primeiro, a detecção automatizada pode refletir vieses (*bias*) dos dados sobre os quais ela é treinada, por exemplo, podendo captar certo tipo de discurso de ódio, mas não para outros para os quais o sistema de detecção não foi “treinado” ou que são “novos”. Esses sistemas também podem ser deliberadamente “enganados”, por exemplo, ao se denunciar postagens não abusivas em alta quantidade ou fugindo delausando palavras obscuras, com mudanças nas ordens das letras ou símbolos. Segundo, as ferramentas automatizadas não são boas em discernir nuances contextuais, ou seja, não fazem exercício interpretativo, por exemplo: quando a nudez deve ser percebida como arte e quando deve ser percebida como obscenidade? Além disso, a participação humana também é importante para se identificar qual será o nível adequado de intervenção (ignorar, exibir um aviso aos espectadores, remover o conteúdo, denunciar à

---

<sup>160</sup> GIANCCHETTA, André Zonaro. *Atuação e responsabilidade dos provedores diante das fake news e da desinformação*. p.27. in *Fake News: a conexão entre a desinformação e o direito*. Coord. Rais, Diogo. São Paulo: Thomson Reuters Brasil, 2018.



polícia, etc). Portanto, algumas dessas plataformas acabam recrutando moderadores humanos. Isso pode ser eficaz para pequenas e médias plataformas, mas as grandes plataformas, como o Facebook ou o Twitter, simplesmente não conseguem monitorar “seus reinos” sem a ajuda de máquinas e inteligência artificial, diante da magnitude e velocidade de dados produzidos por bilhões de usuários diariamente.<sup>161</sup>

Por fim, cumpre deixar claro que não se teve a pretensão aqui de debater os prós e os contras da regulamentação, mas argumentar que alguma forma de regulamentação é urgentemente necessária, para balizar o monitoramento que já acontece pelas plataformas, mas especialmente para garantir transparência e republicanismo dos seus processos internos, maximizando a liberdade de expressão e garantindo outros fins constitucionais. Para isso, a maneira mais eficaz de se reduzir os danos nas redes sociais seria através de uma abordagem baseada em métodos regulatórios alternativos e brandos, existentes em outros setores, que buscam lidar com os riscos de danos naturalmente presentes em suas atividades, por meio da precaução e do cuidado diligente dos agentes regulados.

---

<sup>161</sup> Sobre o tema ver: <[https://motherboard.vice.com/en\\_us/article/xwk9zd/how-facebook-content-moderation-works](https://motherboard.vice.com/en_us/article/xwk9zd/how-facebook-content-moderation-works)> (acessado em 04.02.19).

#### 4. SOBRE A PUBLIC FORUM DOCTRINE: O QUE PENSA A SUPREMA CORTE NORTE AMERICANA SOBRE A LIBERDADE DE EXPRESSÃO EM ESPAÇOS PÚBLICOS: NEUTRALIDADE, RAZOABILIDADE, MANUTENÇÃO DE OUTROS CANAIS DE EXPRESSÃO

De certa forma, as questões subjacentes permeadas a partir de agora envolvem dois temas constitucionais. O primeiro, uma doutrina constitucional incomum e um tanto exótica, baseada na ideia do “*fórum público*” (*public forum doctrine*). A segunda envolve um ideal constitucional geral: a democracia deliberativa. As razões pelas quais se escolheu abordar a *public forum doctrine* ao debater a esfera pública digital residem justamente no fato de que quaisquer regulações que sejam impostas às plataformas no tocante a discursos precisarão lidar com seu embate com o princípio da liberdade de expressão, e na história da Primeira Emenda<sup>162</sup> da Constituição norte-americana e de certa forma também na história da Suprema Corte norte-americana, há *leading cases* que permitiram uma evolução jurisprudencial que acabou por construir balizamentos para regulações na área dos discursos. Usualmente, a questão posta é se o governo tem uma razão legítima e suficientemente forte para restringir o discurso objeto da análise judicial.

Considerando então a ideia geral de que uma sociedade livre cria um conjunto de fóruns públicos, proporcionando o acesso dos indivíduos a uma multiplicidade diversa de pessoas e a uma pluralidade de visões de mundo, e que com isso asseguramos, no processo, que cada um de nós ouça uma ampla gama de opiniões sobre os mais variados assuntos, a concepção de fórum público, aqui analisada a partir do recorte da construção jurisprudencial

---

<sup>162</sup> Primeira Emenda. Constituição dos Estados Unidos. “*Congress shall make no law respecting an establishment of religion, or prohibiting the free exercise thereof; or abridging the freedom of speech, or of the press; or the right of the people peaceably to assemble, and to petition the Government for a redress of grievances.*” Em português: “O Congresso não deverá fazer qualquer lei a respeito de um estabelecimento de religião, ou proibir o seu livre exercício; ou restringindo a liberdade de expressão, ou da imprensa; ou o direito das pessoas de se reunirem pacificamente, e de fazerem pedidos ao governo para que sejam feitas reparações de queixas”. Fonte: <[https://pt.wikipedia.org/wiki/Primeira\\_Emenda\\_%C3%A0\\_Constitui%C3%A7%C3%A3o\\_dos\\_Estados\\_Unidos](https://pt.wikipedia.org/wiki/Primeira_Emenda_%C3%A0_Constitui%C3%A7%C3%A3o_dos_Estados_Unidos)> (acessado em 04.02.19).

da *public forum doctrine* pela Suprema Corte dos Estados Unidos, resumidamente, se utiliza do princípio da liberdade de expressão para proibir um governo (ou o Estado) de “censurar” discursos que ele desaprova.

A doutrina do fórum público é, portanto, uma ferramenta analítica utilizada na jurisprudência da Suprema Corte para determinar a constitucionalidade das restrições de fala implementadas em propriedades públicas. A Corte e os demais Tribunais também empregam a doutrina para decidir se, e entre quais restrições, os indivíduos podem praticar atividades expressivas, como manifestações, peças de teatro, shows, discursos religiosos, etc, em tais propriedades públicas (ou privadas com características públicas, como veremos adiante).<sup>163</sup>

Embora haja discordância sobre a fonte da doutrina, a Suprema Corte identificou repetidamente sua primeira inspiração no caso *Hague v. CIO*, onde o Justice Roberts, encontrando um direito constitucional de usar “ruas e parques para comunicação de opiniões”, baseou-se no fato de que “ruas e parques (...), têm sido imemorialmente mantidos em confiança para o uso do público e, (...) têm sido usados para fins de reunião, comunicação de pensamentos entre os cidadãos e discussões sobre questões públicas.” A Suprema Corte Americana entendeu que as ruas e parques precisam ser mantidos abertos para o público para atividades expressivas, ou seja, para se manifestarem:

Wherever the title of streets and parks may rest, they have immemorially been held in trust for the use of the public and time out

---

<sup>163</sup> Interessantes observações podem ser feitas sobre as propriedades públicas no direito brasileiro. Em primeiro lugar, a própria definição legal está diretamente atrelada ao uso da propriedade, uso este que, por sua vez, dependerá do regime de afetação/desafetação do bem. O artigo 98 do Código Civil conceitua os bens públicos como aqueles pertencentes às pessoas jurídicas de direito público, dividindo estes bens em (i) bens de uso comum do povo, (ii) bens de uso especial e (iii) bens dominicais. Em decorrência de sua afetação, os bens de uso comum do povo e os de uso especial não se submetem, segundo a doutrina clássica, à órbita de aplicação das normas de direito privado, neles aparecendo sensivelmente modificadas as faculdades de uso, gozo e disposição. É importante ter em mente que a derrogação das regras gerais de direito privado (liberdades em geral) para as regras de direito público - as quais, em última análise, ocorrem por meio de restrições/regulamentações - são constitucionalmente justificadas pela persecução de finalidades públicas e constitucionais. Essa observação sobre a classificação dos bens públicos no Brasil quanto ao uso e afetação pode ser importante, porque a incidência de regras de direito público ou de direito privado vai ser decidida justamente a partir da finalidade primária do bem. A conexão com o tema é que, caso enxerguemos as redes sociais menos como editores/imprensa e mais como parques e ruas (bens de uso comum), a incidência de regras de direito público (regulações), são mais facilmente justificáveis.

of mind, have been used for the purposes of assembly, communicating thought between citizens, and discussing public questions. Such use of the streets and public places has, from ancient times, been a part of the privileges, immunities, rights, and liberties of citizens.<sup>164</sup>

No entanto, o direito constitucionalmente garantido de se envolver em protestos públicos não significa livre acesso a todos os tipos de propriedade do Estado. O Estado, não menos do que um proprietário privado, tem o poder de preservar a propriedade sob seu controle para o uso ao qual é legitimamente dedicado. Nada na Constituição dos Estados Unidos - tampouco na brasileira - exige que o governo conceda acesso a todos que desejarem exercer o seu direito à liberdade de expressão em todos os tipos de propriedade, independentemente da natureza da propriedade ou da perturbação que possa ser causada pelas atividades. A Primeira Emenda da Constituição norte-americana nunca significou que as pessoas têm o direito constitucional de protestar, quando, como e onde quiserem, sem qualquer tipo de restrição, aqui enxergada genericamente como regulação.

Kevin O'Neill explica que a Suprema Corte norte-americana sublinhou que no tocante aos assuntos abordados, é axiomático que o Estado não regule o discurso com base em seu conteúdo substantivo ou na mensagem que está sendo transmitida.<sup>165</sup> Ao regular a fala, o Estado não pode favorecer um tipo de discurso em detrimento de outro; a discriminação contra a fala por causa do conteúdo de sua mensagem é considerada inconstitucional. Portanto, quando o Estado não tem como alvo o assunto, porém, ainda mais estreitamente, pontos de vista particulares sobre um determinado tema, a violação aos direitos de liberdade de expressão fica ainda mais evidente. A discriminação de um ponto de vista em específico seria, assim, uma forma notória de discriminação de conteúdo.

Consequentemente, o Estado deve se abster de regular discursos quando a ideologia motivadora específica, a opinião ou a perspectiva do grupo

---

<sup>164</sup> Estados Unidos. Suprema Corte. *Caso Hague v. Committee for Industrial Organization*, 307 U.S. 496 (1939). Tradução livre do autor: "Onde quer que o título de ruas e parques possa descansar, eles foram imemoradamente confiados para o uso do público e do tempo, foram usados para o propósito de reunião, comunicando o pensamento entre os cidadãos e discutindo questões públicas. Esse uso das ruas e lugares públicos tem sido, desde a antiguidade, parte dos privilégios, imunidades, direitos e liberdades dos cidadãos."

<sup>165</sup> O'NEILL, Kevin F. *Disentangling the Law of Public Protest*, 45 *Loyola Law Review* 411 (1999). p. 429.

de falantes seja a justificativa para a imposição de uma restrição, ou seja, quando é a razão (*ratio decidendi*) utilizada para justificar determinada restrição.

Em seu artigo “Disentangling the Law of Public Protest,” O’Neil divide as restrições inadmissíveis com base em conteúdo, originados de julgados da Suprema Corte, em cinco categorias distintas:

- a) [quando] o governo categoricamente suprime ou favorece um tópico ou mensagem em particular - como, por exemplo, no caso *Boos vs. Barry*, onde um estatuto do Distrito de Colúmbia proibiu a exibição de qualquer placa que criticasse um governo estrangeiro a menos de 150 metros de sua embaixada;
- b) [quando] o governo atua como um *gatekeeper* consciente do conteúdo, bloqueando seletivamente o acesso a um fórum baseado na mensagem pretendida pelo interlocutor - como, por exemplo, no caso *Mahoney v. Babbitt*, onde o Serviço Nacional de Parques tentou impedir *banners* de manifestantes antiaborto na rota do desfile inaugural do presidente Clinton;
- c) [quando] o governo submete palestrantes impopulares a uma taxa mais alta por usar um fórum como, por exemplo, no caso *Forsyth County v. Nationalist Movement*, onde, sob um esquema de permissão local, a taxa de proteção policial poderia ser aumentada se o palestrante gerasse controvérsia;
- d) [quando] o governo detém um serviço ou subsídio ao qual o orador teria direito senão por sua mensagem - como, por exemplo, em *Rosenberger v. Rector & Visitors of the University of Virginia*, onde foi negado ao jornal religioso de estudantes o mesmo subsídio para custos de impressão que a universidade forneceu a todas as outras publicações estudantis;
- e) [quando] o governo altera a mensagem pretendida pelo locutor por meio do preço de uso de um fórum - como, por exemplo, no caso *Hurley v. Irish-American Gay, Lesbian & Bisexual Group*, onde, para garantir sua permissão, os organizadores privados do desfile de Saint Patrick’s Day foram compelidos pelo governo a incluir um contingente de manifestantes gays e lésbicas, cuja própria presença transmitia uma mensagem que os organizadores não desejavam transmitir.<sup>166</sup>

---

<sup>166</sup> Tradução livre do autor do trecho: “1) *where the government categorically suppresses or favors a particular topic or message - as, for example, in Boos v. Barry, where a District of Columbia statute banned the display of any sign criticizing a foreign government within 500 feet of its embassy;* 2) *where the government serves as a content-conscious gatekeeper, selectively blocking access to a forum based on the speaker's intended message - as, for example, in Mahoney v. Babbitt,* 3) *where the National Park Service sought to prevent anti-abortion protesters from displaying banners along the route of President Clinton's inaugural parade;* 4) *where the government subjects unpopular speakers to a higher fee for using a forum-as, for example, in Forsyth County v. Nationalist Movement, where, under a local permit scheme, the fee for police protection could be increased if the speaker was likely to generate controversy;* 5) *and where the government alters the speaker's intended message as the price for use of a forum-as, for example, in Hurley v. Irish-American Gay, Lesbian & Bisexual Group,' where, as the price for securing ' their permit, the private organizers of a St. Patrick's Day parade were compelled by the government to include a contingent of gay and lesbian marchers, whose very presence would impart a message that the organizers did not wish to convey.’* In Idem. Ibidem.

Já em relação às restrições impostas ao *modus* (tempo, lugar e forma) da liberdade de expressão em fóruns públicos há, segundo O’Neil, uma tendência na Corte Constitucional de examinar os casos postos a julgamento sob um escrutínio menos rigorosos do que o reservado às regulações baseadas em conteúdo.

Ora, obviamente o Estado pode impor restrições ao tempo, lugar e maneira de manifestações em lugares públicos. A questão é como o faz e se o faz baseado em razões legítimas ou não. Qualquer ponderação de interesses minimamente razoável leva à conclusão que ninguém tem o direito de usar alto-falantes e carros de som nas ruas às três horas da manhã, a fim de se protestar sobre a taxa de homicídios, crimes, racismo, mudança climática ou qualquer outro tema mais ou menos polêmico. No entanto, as restrições de tempo, lugar e forma devem ser razoáveis e limitadas, de modo que o Estado é essencialmente obrigado a permitir que os cidadãos, quaisquer que sejam suas visões de mundo, usem fóruns públicos para transmitir mensagens de sua escolha.

Ao descrever sobre as formas de regular o tempo, lugar e a forma de se expressar, Kevin O’Neil explica que tais regulamentações são frequentemente mantidas e representam uma parte comum do cenário regulatório na maioria das cidades. A análise de tempo, lugar e forma, segundo o autor, “*é governada por um teste de três partes, exigindo uma investigação separada em três questões distintas: (1) o regulamento é verdadeiramente neutro em termos de conteúdo; (2) se o regulamento é estreitamente adaptado para servir a um interesse governamental significativo; e (3) se o regulamento deixa em aberto canais alternativos de comunicação suficientes.*”<sup>167</sup>

Nesse sentido, quanto à primeira parte do teste, o principal questionamento para determinar a neutralidade de conteúdo é saber se o governo restringiu este ou aquele discurso em específico por discordar da mensagem transmitida. Neste caso, o motivo utilizado para justificar a regulação (*ratio*) precisa ser o propósito ou a intenção do Estado para

---

<sup>167</sup> Tradução livre do autor do trecho “is governed by a three-part test, requiring separate inquiry into three distinct issues: (1) whether the regulation is truly content neutral; (2) whether the regulation is narrowly tailored to serve a significant government interest; and (3) whether the regulation leaves open sufficient alternative channels of communication.” In Id. p. 433-435.

estabelecer a restrição. Assim, uma restrição que sirva a propósitos não relacionados ao conteúdo da expressão é considerada neutra, mesmo que tenha um efeito acidental em um grupo determinado de falantes ou mensagens, mas não em outros. Em suma, a regulação governamental de atividade expressiva será considerada neutra, desde que sua justificativa não faça referência ao conteúdo do discurso regulado.

O autor exemplifica casos em que as restrições de fala foram consideradas neutras em termos de conteúdo, mesmo atingindo de maneira mais severa um orador ou mensagem em particular, desde que o Estado tenha justificado as restrições para fins que não tenham nada a ver com o conteúdo do discurso:

(1) uma restrição de ruído limitando o nível de decibéis para concertos no Central Park - embora a restrição se mostrasse especialmente onerosa para os músicos de rock - onde o propósito declarado do governo era preservar a quietude da propriedade adjacente;

(2) [uma restrição] em acampar no Lafayette Park e no Shopping - apesar de sua aplicação contra os moradores de rua os impedirem de dormir durante a noite em tendas perto da Casa Branca - onde o propósito da proibição era manter os parques de Washington em uma condição atraente e intacta;

(3) [uma restrição] que proíbe a manutenção durante a noite de quaisquer 'adereços' no território do Capitólio dos EUA - apesar de efetivamente frustrar um plano de defensores dos sem-teto de erigir, como parte de uma vigília de sete dias, uma estátua de barro de 500 pounds de um homem, mulher e criança se amontoando sobre uma grelha a vapor - onde a proibição da noite era justificada como garantia do controle diário e significativo do governo sobre os terrenos do Capitólio;

e (4) uma ordem proibindo expressivamente a atividade dentro de um local da Floresta Nacional de San Juan fechada para atividade madeireira - embora o impacto da proibição tenha sido unilateral, já que apenas ambientalistas tentaram protestar entre as árvores - onde o governo justificou sua proibição de proteger 'saúde e segurança' e ... propriedade'.<sup>168</sup>

---

<sup>168</sup> Tradução livre do autor do trecho "(1) a noise regulation limiting the decibel level at Central Park concerts-even though the restriction proved especially burdensome for rock musicians - where the government's stated purpose was to preserve the quietude of adjacent property; (2) a National Park Service ban on camping in Lafayette Park and the Mall-even though its enforcement against homeless advocates prevented them from sleeping overnight in "tent cities" near the White House-where the ban's underlying purpose was to maintain Washington's parks in an attractive and intact condition; (3) a regulation banning the overnight maintenance of any "props" on the U.S. Capitol grounds-even though it effectively thwarted a plan by homeless advocates to erect, as part of a seven-day vigil, a 500-pound clay statue of a man, woman, and child huddled over a steam grate-where the overnight ban was justified as affording the government meaningful day-to-day control over the Capitol grounds; and (4) an order banning all expressive activity within a sector of the San Juan National Forest closed for logging activity-even though the ban's impact was entirely one-sided, since only environmentalists sought to protest among the trees-where the government justified its ban as protecting health and safety and... property."Id.

Já em relação à ponderação de interesses em jogo, ou seja, se a restrição era estritamente adequada para servir a um interesse estatal significativo, algo bastante parecido à fase da adequação no teste da proporcionalidade, para ser sustentado sob a Primeira Emenda, uma restrição de *tempo, lugar e forma* deve ser estritamente adaptada para servir a um interesse significativo. Embora este requisito seja tecnicamente composto por dois componentes distintos; a uma, deve ser estritamente adequado, e a duas, servir a um interesse público significativo, este último fator raramente se mostrou relevante na jurisprudência da Corte, afinal, é muito complicado a identificação do que seria um interesse governamental "significativo".

Invariavelmente, a análise se volta para a presença ou ausência de adaptação estreita da medida restritiva (o que também se aproxima bastante do teste da proporcionalidade em sentido estrito). Nesse sentido, O'Neil lembra que a Suprema Corte estabeleceu que as restrições de *tempo, lugar e forma* precisam ser necessariamente as menos restritivas ou menos impactantes para atingir aos fins buscados. Assim, regulações sobre atividades expressivas que impõem, por exemplo, proibições totais de desfiles, manifestações, greves, panfletagem de porta em porta ou panfletagem pública, usualmente falham a esta etapa do teste.<sup>169</sup>

Por último, para ser válida de acordo com a Primeira Emenda, uma restrição de *tempo, lugar e forma* deve deixar outros canais amplamente abertos para a comunicação da mensagem dos cidadãos. Não obstante uma regulação do discurso possa falhar com este requisito caso impeça formas de expressão muito menos dispendiosas do que as alternativas viáveis, o teste básico para medir a suficiência de canais alternativos constitui na averiguação da disponibilidade de um fórum público acessível e onde se espera que o público-alvo da manifestação pretenda passar. Senão não há nenhum sentido em haver a manifestação.

Ao realizar esta análise, um Tribunal deve considerar, segundo O'Neil, "(1) a audiência pretendida pelo falante e (2) a extensão em que o local escolhido contribui para sua mensagem." A conclusão é que uma restrição não

---

<sup>169</sup>Id.p 438-440.



deixa abertos canais alternativos amplos se a mensagem não for capaz de alcançar seu público pretendido.

Além disso, a restrição pode ser considerada inválida se priva os comunicadores de um modo de comunicação excepcionalmente valioso ou importante, ou se ameaça a sua capacidade de se comunicar efetivamente.<sup>170</sup> É especialmente neste ponto que as restrições à liberdade de expressão eventualmente impostas em plataformas interativas - caso entendidas como fóruns públicos - será mais saliente: impedir esse ou aquele tipo de discurso em uma grande rede social mantém outros canais efetivos abertos? Ou seja, as redes sociais são comunicacionalmente tão valiosas que impedir o acesso a elas seria uma restrição inconstitucional?

Nos casos *Dr. Martin Luther King, Jr. Movement*,<sup>171</sup> *Students Against Apartheid*,<sup>172</sup> and *Bay Area Peace Navy*<sup>173</sup> as restrições impostas não forneceram canais alternativos suficientes, merecendo ser comentados. Na primeira hipótese, uma organização de direitos civis negros visava marchar em um bairro predominantemente branco, mas as autoridades da cidade negaram a permissão para uma segunda marcha pelo mesmo bairro. Sua primeira tentativa havia sido frustrada por moradores e espectadores que atiraram pedras, tijolos e dispositivos explosivos nos manifestantes, de forma que as autoridades da cidade negaram a permissão para uma segunda marcha pelo mesmo bairro, propondo, em vez disso, uma rota alternativa através de um bairro totalmente negro.

Ora, uma vez que o objetivo da marcha era justamente divulgar e protestar contra um padrão de violência contra os negros que tentavam residir ou viajar pelo bairro específico, a Corte sustentou que a imposição por um caminho alternativo, ou seja, que afastava a manifestação do bairro objeto do discurso e distanciava a mensagem do público-alvo pretendido, era constitucionalmente inadequada como um canal alternativo de comunicação.

---

<sup>170</sup> Neste ponto, proibir o acesso discursos em redes sociais poderia ser interpretado como uma restrição inconstitucional.

<sup>171</sup> Estado Unidos. Suprema Corte. *Dr. Martin Luther King, Jr. Movement, Inc. v. City of Chicago*, 419 F. Supp. 667, 672 (N.D. 11. 1976).

<sup>172</sup> Estados Unidos. Suprema Corte. *Students Against Apartheid Coalition v. O'Neil*, 660 F. Supp. 333, 335-37 (W.D. Va. 1987).

<sup>173</sup> Estados Unidos. Suprema Corte. *Bay Area Peace Navy v. United States*, 914 F.2d 1224, 1225-26, 1231 (9th Cir.1990).

Em *Students Against Apartheid Coalition v. O'Neil*, o sucesso das manifestações desafiou regulamentos internos da Universidade da Virgínia sobre o uso da grama e do *campi*, sob os quais eles foram impedidos de montar barracas para protestar contra o apartheid sul-africano e forçar o corpo diretivo da Universidade a adotar uma política de desinvestimento em relação à África do Sul. O público-alvo do discurso era o corpo diretivo da Universidade, cujas reuniões eram realizadas em frente ao local onde os manifestantes pretendiam erguer objetos de protesto e montar barracas. No caso, as restrições impostas pela Universidade permitiria a exposição dos objetos apenas em áreas além do alcance da visão (ou da visão clara) do local exato onde os diretores se reuniam. Neste caso, a Corte considerou que tal solução não oferecia canais alternativos de comunicação, portanto, seria inconstitucional.

No *leading case Bay Area Peace Navy v. United States*, o Nono Circuito declarou que a criação de uma zona de segurança de setenta e cinco jardas (cerca de 70 metros) estabelecida em torno do mirante à embarcações, na hipótese de um desfile naval, violavam os direitos da Primeira Emenda de manifestantes que esperavam apresentar mensagens anti-guerra com *banners* e outros objetos de manifestação. O Tribunal concluiu que as alternativas sugeridas pela administração (distribuir panfletos ou protestar em terra) não poderiam substituir uma procissão pela água - porque o público-alvo pretendido pela conteúdo da manifestação não restou acessível.

Em contraste com as decisões anteriores, há decisões da Corte e outros Tribunais em que as restrições de fala impostas foram mantidas como hipóteses em que foram proporcionados canais alternativos de comunicação suficientemente aceitáveis, como nos casos *Taxpayers for Vincent*,<sup>174</sup> *ACORN*,<sup>175</sup> e *Fee*.<sup>176</sup>

No primeiro caso, um candidato político contestou, sem sucesso, uma portaria que proibia a disposição de cartazes em propriedades públicas e,

<sup>174</sup> Estado Unidos. Suprema Corte. *City Council of Los Angeles v. Taxpayers for Vincent*, 466 U.S. 789, 817 (1984);

<sup>175</sup> Estados Unidos. Corte do Nono Circuito. *ACORN v. City of Phoenix*, 798 F.2d 1260, 1273 (9th Cir. 1986);

<sup>176</sup> Estados Unidos. Suprema Corte. *United States v. Fee*, 787 F.Supp. 963, 969-70 (D. Colo. 1992).

assim, impedia a prática adotada pelo candidato de prender objetos de campanha à fiação pública de luz. Argumentando que a propriedade coberta pela portaria (que incluía postes de iluminação, luminárias, hidrantes e troncos de árvores) não era um fórum público, a Corte determinou que o Decreto ultrapassava o teste de razoabilidade como uma restrição de conteúdo neutro à desordem visual. Ao chegar a essa ponderação, a Corte concluiu que outros canais alternativos de comunicação estavam prontamente disponíveis ao demandante; ele permanecia livre para falar e distribuir panfletos precisamente nos mesmos locais públicos onde pendurar sinais era proibido - apenas não nos fios de luz -, e por entender que nada indicava que a manifestação nos fios de luz era excepcionalmente superior a outros modos alternativos de expressão.

Na caso *ACORN*, o Nono Circuito manteve uma portaria que proibia a solicitação de contribuição pecuniária de motoristas parados em cruzamentos. O Tribunal concluiu que a restrição deixava aberto amplos caminhos alternativos para a comunicação: os manifestantes eram livres não apenas para usar outros métodos de captação de recursos diversos, empregados por outras organizações (incluindo solicitação de contribuição nas calçadas, porta-a-porta, e mala direta), mas também estavam livres para abordar os veículos parados com panfletos, desde que não pedissem dinheiro.

Em *Fee*, nove manifestantes ambientais foram condenados por desafiar intencionalmente uma ordem especial do Serviço Florestal que fechou um setor da Floresta Nacional de San Juan para a atividade madeireira. Os réus entraram no setor fechado da floresta com o intuito de protestar contra o corte de árvores. O Tribunal entendeu que a ordem que proibia toda atividade expressiva dentro do local durante um período de noventa dias como um *tempo, lugar e restrição de forma* como sendo razoável. A ordem havia sido precedida por protestos anteriores em que os manifestantes se acorrentaram a uma árvore por doze dias, prenderam-se a outros objetos no local, bloquearam veículos, rolaram troncos para a estrada e cercaram madeireiros e árvores para impedir o corte. Não obstante quisessem protestar onde as árvores estavam de fato em perigo, os manifestantes tinham permissão para ficar dentro da floresta, e exatamente no ponto em que a estrada entrava na área de extração de madeira. Isso, segundo a Corte, era constitucionalmente suficiente. Ao

manter a ordem do Parque, concluiu-se que a restrição manteve abertos canais alternativos suficientes pelos quais os manifestantes poderiam comunicar sua mensagem.<sup>177</sup>

Ainda na obra em comento, Kevin O'Neil trata de casos envolvendo a noção de fóruns públicos privados, ou seja, propriedades privadas que resguardam características análogas aos fóruns públicos tradicionais - a exemplo dos Shoppings recorrentemente discutidos nas Cortes Superiores e nos Tribunais locais. É interessante notar que em um *leading case*<sup>178</sup> envolvendo Shopping em 1972, a Suprema Corte norte-americana entendeu que a Primeira Emenda serviria apenas para impedir restrições Estatais aos discursos em propriedades públicas, não se estendendo aos *Shopping-Malls*.

No entanto, a partir desse precedente firmado, advogados começaram a invocar argumentos alternativos baseados em cláusulas de liberdade de expressão contidas nas Constituições Estaduais. Essas cláusulas serviram como argumentos viáveis, pois ao construir entendimentos sobre suas próprias constituições, juízes estaduais estariam livres para ampliar o espectro protecional de liberdades individuais em comparação aos entendimentos firmados em sede federal - com base na Constituição dos Estados Unidos -, mesmo em casos em que os textos constitucionais federais e estaduais eram similares.<sup>179</sup> As disposições sobre liberdade de expressão das constituições estaduais podem ser agrupadas em três categorias:

- (1) aquelas que emulam a linguagem exclusivamente negativa da Primeira Emenda federal (isto é, "o Congresso não fará nenhuma lei ..."); (2) aquelas que conferem direitos de fala em ampla linguagem afirmativa; e (3) aquelas que combinam cláusulas positivas e negativas. Essas expressões afirmativas, associadas ou não a uma cláusula negativa, fornecem um notável contraste com a carta federal.<sup>180</sup>

<sup>177</sup> O'NEILL, Kevin F. *Disentangling the Law of Public*, op. cit. nota 166. p. 446-447.

<sup>178</sup> Estados Unidos. Suprema Corte. *Lloyd Corp. v. Tanner*, 407 U.S. 551, 567 (1972).

<sup>179</sup> Estados Unidos. Suprema Corte. *City of Mesquite v. Aladdin's Castle, Inc.*, 455 U.S. 283, 293 (1982); e *PruneYard Shopping Ctr. v. Robins*, 447 U.S. 74, 81 (1980).

<sup>180</sup> O'NEILL, Kevin F. *Disentangling the Law of Public*, op cit. nota 166. p. 455. Tradução livre do autor do trecho "State free speech provisions may be grouped into three categories: (1) those that emulate the exclusively negative language of the federal First Amendment (i.e., "Congress shall make no law . . . ,"); (2) those that confer speech rights in sweeping affirmative language; and (3) those that combine affirmative and negative clauses. These affirmative expressions, whether or not coupled with a negative clause, provide a striking contrast to the federal charter. Pennsylvania's provision is a good example: "The free communication of thoughts and opinions is one of the invaluable rights of man, and every citizen may freely speak, write, and print on

O'Neil, ainda no ano de 1999, quando a Internet e as tecnologias a ela associadas não possuíam a centralidade na vida humana que detém hoje, já havia alertado que

[a]o final do século, nossos fóruns públicos tradicionais (calçadas, parques e praças públicas), onde o discurso desfruta de maior proteção, são cada vez menos as encruzilhadas da comunidade, cada vez menos o cenário em que encontramos nossos concidadãos a pé. Enquanto isso, nosso dia-a-dia é cada vez mais realizado em novos ambientes para espaços de tráfego de pedestres que são de propriedade privada e, portanto, não oferecem acesso constitucional federal à atividade expressiva. À medida que a tendência para a privatização do espaço "público" continuar, a batalha pelo acesso à liberdade de expressão nesse espaço continuará também - e, como a Primeira Emenda não tem aplicação nesta esfera privada, a batalha será travada inteiramente por uma base de Estado por Estado.<sup>181</sup>

Apresentados brevemente alguns dos principais casos da jurisprudência norte-americana sobre a *public forum doctrine*, e considerando que a finalidade deste tópico está em estabelecer possíveis interseções entre as restrições governamentais impostas à liberdade de expressão na esfera pública clássica e eventuais restrições na esfera pública digital, devemos adentrar também o debate sobre o fato das plataformas interativas serem (ou se deveriam passar a ser) reconhecidas como fóruns públicos, pelo menos para fins de regulação dos discursos. O objetivo, portanto, é delinear os requisitos constitucionais genéricos mínimos para que uma restrição seja constitucionalmente válida, tendo em vista o fato de que quaisquer regulações que procurem prevenir *online harms* também acabam esbarrando na liberdade de expressão. Antes disso, no entanto, é crucial diferenciar características do ambiente informacional no qual a jurisprudência sobre a Primeira Emenda norte-

---

*any subject, being responsible for the abuse of that right." Absent from this language is any suggestion that its sole purpose is to restrain the government. Thus, affirmative phrasing in a state free speech clause raises the possibility that it reaches not merely governmental but also private regulation of speech.*" Ver também Constituição da Pennsylvania. art. 1, § 7. Linguagem virtualmente idêntica pode ser encontrada no artigo II, seção 6 da Constituição de Arkansas e art. I, seção 19 da Constituição do Tennessee.

<sup>181</sup>O'NEILL, Kevin F. *Disentangling the Law of Public*, op.cit. nota 166. p 456. Tradução livre do autor do trecho "At the century's end, our traditional public fora (sidewalks, parks, and public squares) where speech enjoys the greatest protection, are less and less the crossroads of the community, less and less the setting where we encounter our fellow citizens on foot. Meanwhile, our day-to-day lives are increasingly carried out in new environments for pedestrian traffic-spaces that are privately owned and thus afford no federal constitutional access for speech activity. As the trend toward privatization of "public" space continues, the battle for speech-related access to that space will continue as well-and, since the First Amendment has no application in this private sphere, the battle will be waged entirely on a state-by-state basis."

americana evoluiu, das características atuais do ecossistema informacional das Sociedades Algorítmicas, aprofundadas no primeiro e segundo capítulos do trabalho.

Parece haver boas razões para expandir a noção de fórum público para além dos lugares clássicos como ruas, parques, e prédios públicos, pois, na era moderna, outros lugares vêm ocupando cada vez mais o papel de fóruns públicos tradicionais. A grande mídia e a Internet tornaram-se muito mais importantes do que ruas e parques, no sentido de locus expressivo. Não obstante, nota-se que a Suprema Corte tem sido cautelosa em expandir a doutrina do fórum público para além das ruas e parques. A primeira razão pode ser que qualquer expansão da doutrina esbarra no fato de que as intermediárias da Internet são legalmente instituições privadas, que não estão totalmente cobertas pela Primeira Emenda. Portanto, se o Facebook ou o Twitter recusarem-se a publicar certos discursos ou retirar alguns das plataformas, desde de que fundamentado em razões minimamente razoáveis - o que recorrentemente fazem com base em suas regras de comunidade -, segundo uma leitura restrita da jurisprudência, não haveria problema constitucional, pois tais lugares seriam tecnicamente propriedades privadas. Uma segunda possível razão decorre da crença de que, uma vez abandonada a jurisprudência clássica, as linhas entre público e privado serão extremamente difíceis de serem traçadas, e o Judiciário pode vir a ser recorrentemente demandado a decidir sobre a possibilidade de acesso às plataformas por este ou aquele discurso. Sendo assim, nos deparamos aqui com o sério problema de como “transportar ”a idéia do fórum público para o moderno ambiente tecnológico.

Neste ponto, é importante observar que caso uma hipotética mudança na concepção das redes sociais não venha pela via Judiciária, é possível que tal transposição também seja feita por outros atores, em conjunto ou não. O Executivo, o Legislativo, as instituições privadas e os cidadãos comuns podem considerar fazer exatamente isso. Em outras palavras, embora as Cortes constitucionais talvez não estejam preparadas para dizer que a idéia do fórum público se estende além das ruas e parques como uma questão de direito constitucional, os outros Poderes e as próprias plataformas podem concluir que uma sociedade livre requer o direito de acesso a locais onde as pessoas se

encontram e debatem ideias. Ou seja, não só leis emanadas do Legislativo podem conceber as plataformas interativas como fóruns públicos, quanto hipotéticas regulamentações do Executivo poderiam aplicar restrições de direito público constitucionalmente justificadas, bem como as próprias empresas de tecnologia poderiam fundamentar essa ou aquela restrição no fato de que entendem seu reino como um fórum público.

Em resumo, vimos que os fóruns públicos, na esfera pública clássica, desempenham papéis importantes na facilitação da liberdade de expressão, especialmente no que tange aos discursos políticos, e que para restringir o direito à fala em tais locais, o Estado precisa de uma justificativa legítima e satisfatória. Vimos também que as regulações aos discursos precisam ser neutras, pois não podem estar embasadas no conteúdo da mensagem em si, havendo também a necessidade de serem mantidos abertos outros canais de comunicação. Nota-se, entretanto, que a jurisprudência sobre a Primeira Emenda norte-americana iniciou-se na década de 1920, quando a principal ameaça ao ambiente dos discursos políticos nos Estados Unidos era a supressão e perseguição de dissidentes pelo Estado. Temos, assim, que a jurisprudência construiu-se a partir dos desafios comunicacionais da época, alguns ainda existentes - como a perseguição a opositores; outros, já ultrapassados.

Tim Wu, professor de Direito da Universidade de Columbia, explica que a jurisprudência moldada a partir do início do século XX pressupõe um mundo pobre em informação e se concentra exclusivamente na proteção de indivíduos do Estado “*como se fossem borboletas raras e delicadas, ameaçadas por um monstro terrível.*”<sup>182</sup> Acontece que, como explicado em tópico anterior, o *overload* informacional, causado pelo fim das barreiras de publicação, tornou a informação abundante, produzida por “amadores” que influenciam o debate político pelas redes sociais, dentro de um contexto de filtros bolhas, cascatas cibernéticas e de manipulação de *trending topics* por contas automatizadas (*bots*). Sendo assim, em contraste com momentos históricos anteriores - protagonizados pela mídia tradicional, cujos padrões ético-jornalísticos e de *gatekeeping* eram minimamente aceitáveis -, o problema agora enfrentado é

---

<sup>182</sup> WU, Tim. *Is the First Amendment Obsolete?* In. op. cit. nota 30. p. 2. Disponível em <https://ssrn.com/abstract=3096337>

outro. Mantendo a analogia proposta por Wu: as borboletas raras e delicadas, agora, estão mais parecidas com mariposas; a oferta é aparentemente infinita:

Em outras palavras, se já foi difícil falar, agora é difícil ser ouvido. Dito de outra forma, não é mais a fala ou a informação que é escassa, mas a atenção dos ouvintes. Ao contrário da década de 1920, a informação é abundante e falar é fácil, enquanto o tempo e a atenção do ouvinte se tornaram commodities altamente valorizadas.<sup>183</sup>

Infelizmente, no contexto das Sociedades Algorítmicas, a liberdade de expressão sem freios e restrições - ou melhor, o *overload* de discursos na esfera pública digital - pode ser usado para atacar, assediar e silenciar tanto quanto é usado para iluminar ou debater. E o uso da fala como uma ferramenta para suprimir a fala é, por sua natureza, algo muito novo e desafiador para a jurisprudência sobre liberdade de expressão. Diante de tais desafios, aplicar friamente a *public forum doctrine* e outras doutrinas da liberdade de expressão, construídas sob um contexto informacional distinto, parece, na melhor das hipóteses, uma ideia ultrapassada. No entanto, considerando que essas doutrinas podem na verdade acabar bloqueando os esforços para lidar com algumas das causas da desinformação, é crucial reajustá-las ao contexto informacional das Sociedades Algorítmicas.

Jack M. Balkin, já havia alertado em outra publicação<sup>184</sup> que a era digital iria alterar a importância dos direitos de liberdade de expressão, e, afirmado que a Primeira Emenda da Constituição norte-americana seria inadequada para atingir os objetivos principais do princípio da liberdade de expressão, ou seja, a promoção e a proteção de uma cultura democrática.

É bom lembrar que nos estágios iniciais da Internet, a *World Wide Web*, foi festejada como uma ferramenta capaz de amplificar a democracia e a participação popular. Nas palavras de Balkin:

A Internet inicial parecia simbolizar as possibilidades de uma cultura democrática. Agora, as pessoas poderiam se tornar suas próprias emissoras, falando com um número indefinido de pessoas, não apenas em seus próprios países, mas em todo o mundo. Armados

---

<sup>183</sup> Op. cit. p. 7. Tradução livre do autor para o trecho "In other words, if it was once hard to speak, it is now hard to be heard. Stated differently, it is no longer speech or information that is scarce, but the attention of listeners. Unlike in the 1920s, information is abundant and speaking is easy, while listener time and attention have become highly valued commodities."

<sup>184</sup> BALKIN, Jack M. *Cultural Democracy and the First Amendment*. 110 Nw. U. L. Rev. 1053 (2016). Disponível em <<https://scholarlycommons.law.northwestern.edu/nulr/vol110/iss5/3>> (acessado em 04.02.19).



com tecnologias digitais, os indivíduos comuns poderiam encontrar rotas ao redor dos *gatekeepers* da mídia tradicional. Eles também poderiam participar da cultura de maneiras sempre novas através de novas formas digitais de produção e apropriação cultural.<sup>185</sup>

Entretanto, com o exponencial desenvolvimento da Internet e das tecnologias a ela associadas, Balkin argumenta que, caso os Tribunais continuem interpretando a Constituição restritivamente, a Primeira Emenda - a qual aqui podemos estender para o entendimento moderno de liberdade de expressão, genericamente falando - se mostra ainda mais inadequada para promover esses objetivos democráticos. Assim, o desenvolvimento de doutrinas jurisprudenciais sobre liberdade de expressão, na verdade, acaba dificultando a proteção e o desenvolvimento de uma cultura verdadeiramente democrática:

Em primeiro lugar, argumentei que, à medida que a Internet se desenvolvesse, as doutrinas da primeira emenda feitas pelos juízes, embora necessárias à proteção de uma cultura democrática, seriam cada vez mais insuficientes. Grande parte da responsabilidade de garantir uma cultura democrática caberia a atores privados, tecnólogos, legisladores e agências administrativas. Em segundo lugar, argumentei que à medida que os modelos de negócios se desenvolvem, as empresas empregam cada vez mais teorias da Primeira Emenda não apenas para proteger a liberdade de expressão em geral, mas também para restringir o acesso à infraestrutura digital e impedir a regulamentação para promover seus modelos de negócios e proteger seus lucros.<sup>186</sup>

---

<sup>185</sup> Tradução livre do autor para o trecho “*The early Internet seemed to symbolize the possibilities for such a democratic culture. Now people could become their own broadcasters, speaking to an indefinite number of people, not only in their own countries, but around the world. Armed with digital technologies, ordinary individuals could route around traditional media gatekeepers. They could also participate in culture in ever new ways through new digital forms of cultural production and appropriation.*” in BALKIN, Jack M. *Free Speech in the Algorithmic Society: Bid Data, Private Governance, and New School Speech Regulation*, opt. cit. nota 8. p. 5.

<sup>186</sup> Idem. Ibidem. p. 2.

## 4.1 Devemos enxergar as redes sociais como fóruns públicos?

Em 2017, a Suprema Corte norte-americana já havia expressamente reconhecido no caso *Packingham v. North Carolina* a importância das redes sociais para a deliberação na esfera pública, quando dispôs que

[w]hile in the past there may have been difficulty in identifying the most important places (in a spatial sense) for the exchange of views, today the answer is clear. It is cyberspace...and social media in particular.<sup>187</sup>

A essa altura da vida conectada, parece que essa afirmação encontra respaldo numa simples observação da realidade mundana. Inclusive, essa proposição parece ser quase que axiomática: não há mais qualquer dúvida de que as plataformas interativas representam o locus no qual as pessoas deliberam.

Porém, embora exigir das Cortes Constitucionais a criação de entendimentos nesse sentido possa parecer um pedido para se atestar o óbvio, o que se tem em mente, na verdade, é que, a partir do momento em que as discussões jurídicas sobre censura (governamental e privada) passarem a ser recorrentes no Judiciário, e com elas os processos judiciais que demandem análises do contexto das situações de fala e das funções desempenhadas pela rede social e pelos envolvidos nos casos concretos, o que parecia ser o atestado do óbvio, pode se mostrar extremamente complicado.

Se o Judiciário já vem sendo rotineiramente demandado em ações indenizatórias por conteúdo postado em rede social, em um futuro próximo, caso continuem crescendo as demandas por transparência na governança das plataformas e pela regulação dos discursos nas grandes redes sociais, ao que tudo indica, os juízes serão constantemente demandados a decidirem sobre a legalidade e a constitucionalidade de restrições à liberdade de expressão, de modo que a análise do fórum e das restrições possíveis possa vir a se tornar

---

<sup>187</sup> Tradução livre do autor de trecho constante no caso *Packingham v. North Carolina*: [e]mbora no passado possa ter havido dificuldade em identificar os lugares mais importantes (num sentido espacial) para a troca de pontos de vista, hoje a resposta é clara. É o ciberespaço... e as redes sociais em particular.

uma das ponderações mais comuns nos processos judiciais que versem sobre matérias como monitoramento, censura e aplicabilidade dos termos de uso.

Embora a Corte tenha notado que as redes sociais são, atualmente, os lugares mais importantes para a troca de ideias, não podemos simplesmente concebê-las como lugares públicos físicos, e muito menos aplicar a lei sem a devida adequação. Essa transposição não pode ser direta, porque, em primeiro lugar, tecnicamente as redes sociais são lugares privados, e, em segundo, porque a própria “*forum analysis*” é complexa, existindo uma variedade de fóruns públicos na jurisprudência da Suprema Corte norte-americana, cada qual com seu grau de restringibilidade. Sendo assim, não se está aqui a dizer que as redes sociais são fóruns públicos, mas que, em certas ocasiões, devem ser enxergadas como tal.

Por esse motivo, a decisão judicial da *United States District Court for the Southern District of New York* no caso *Knight First Amendment Institute v. Trump* representa um marco nas discussões sobre regulação da liberdade de expressão na esfera pública digital, pois aplica a mesma ferramenta analítica acima explicada, desenvolvida inicialmente para espaços físicos reais, para proibir um Presidente de bloquear usuários de sua conta no Twitter, ou seja, na esfera digital. Antes de comentarmos a decisão, entretanto, vale resumir como a Suprema Corte norte-americana realiza a análise do fórum no qual as manifestações expressivas ocorrem. Em 1983, no caso *Perry Educ. Ass’n v. Perry Educators’ Ass’n*,<sup>188</sup> a Suprema Corte dividiu os fóruns em três tipos: (i) fóruns públicos tradicionais, (ii) fóruns públicos designados e (iii) fóruns não públicos.

Os fóruns públicos tradicionais incluem parques públicos, calçadas e áreas que tradicionalmente são abertas a discursos e debates políticos, havendo uma forte carga histórica de expressividade no local. Nessas áreas os manifestantes desfrutam das mais fortes proteções da Primeira Emenda. Portanto, nos fóruns públicos tradicionais, o Estado não pode discriminar os discursos com base nas opiniões dos oradores, pois ao fazê-lo estaria discriminando um ponto de vista específico (*viewpoint discrimination*), o que é proibido pela Primeira Emenda. No entanto, como explicado ao longo deste

---

<sup>188</sup> Informação disponível em: <<https://supreme.justia.com/cases/federal/us/460/37/>> (acessado em 04.02.2019).

trabalho, o Estado pode sujeitar a expressividade a restrições razoáveis, de conteúdo neutro, em seu tempo, lugar e forma. Assim, ao considerar as restrições de discurso nos fóruns públicos tradicionais, os Tribunais elevam o grau de escrutínio, de modo que só são permitidas restrições que legitimamente e justificadamente atendem a um interesse público ou estatal e que são estreitamente adaptadas para atender às necessidades desse interesse.

Já os fóruns públicos designados incluem teatros municipais e salas de reunião em universidades.<sup>189</sup> Nesse sentido, os Tribunais não classificam um fórum como sendo um fórum designado sem que tenha havido uma indicação clara da intenção pública em abrir o fórum. Tal intenção deve ser averiguada em parte com base em "*policy and practice*", ou seja, o que tem sido costumeiramente feito nesses locais, e se a propriedade é de um tipo compatível com a atividade expressiva pretendida. Em suma, são lugares que às vezes estão abertos para a expressão pública, mesmo que a propriedade pública não seja fórum público tradicional. A questão é: uma vez aberto um fórum público designado, embora o Estado não seja obrigado a mantê-lo aberto, enquanto o mantiver o discurso praticado nos fóruns designados recebem as mesmas proteções de Primeira Emenda dadas aos fóruns públicos tradicionais.

Dentro dos fóruns públicos designados existe uma subdivisão chamada de fórum limitado, ou seja, é um tipo específico de fórum público designado. Nestes casos, o Estado pode limitar o acesso a um fórum público para certas classes ou tipos de discurso. Laryssa Lindsy explica que essa é a categoria com maiores divergências jurisprudenciais:

A categoria mais difícil, ou subcategoria, é o fórum público limitado. Doutrinariamente, é aqui que as coisas começam a ficar confusas.

---

<sup>189</sup> "The designated public forum "consists of public property which the state has opened for use by the public as a place for expressive activity." Courts will not find a designated public forum absent a clear indication of government intent to open the forum, though such intent can be determined in part based on "policy and practice" and whether the property is of a type compatible with expressive activity. Examples of designated public forums include a university-created "campus free speech zone" open to all speakers or meeting rooms available for use by any member of the public, such as a public library." in Lidsky, Lyrissa B. *Government Sponsored Social Media and Public Forum Doctrine under the First Amendment: Perils and Pitfalls*, 19 Pub.Law. 2 (2011). Disponível em: <<http://scholarship.law.ufl.edu/facultypub/626>> (acessado em 04.02.19)

Independentemente da categorização, o fórum público limitado é um lugar ou espaço designado para a fala por 'certos grupos' ou para 'a discussão de certos assuntos'. Por exemplo, uma Universidade pode limitar um fórum público que estabelece uso por grupos de estudantes, e um distrito escolar pode limitar um fórum público à discussão de 'negócios do conselho escolar'. (...) Além disso, as decisões da Suprema Corte nessa área são quase sempre decididas por 5 a 4 votos, tornando ainda mais difícil prever quanta liberdade de governo tem que exercer controle editorial em um fórum público limitado.<sup>190</sup>

No entanto, caso haja a restrição a uma atividade expressiva ou a um orador em específico, o qual hipoteticamente se enquadra na classe para a qual o fórum público estaria geralmente disponível, a restrição estará novamente sujeita ao escrutínio rigoroso dos fóruns tradicionais. Por conseguinte, o estabelecimento e a aplicação de restrições no fórum público limitado devem ser neutros e razoáveis à luz dos objetivos servidos pelo fórum.

Por fim e por exclusão, os fóruns não públicos são fóruns de discurso público que não são fóruns públicos tradicionais nem fóruns públicos designados. De acordo com a Suprema Corte, em um fórum não público, o Estado pode restringir o conteúdo de um discurso, desde que a restrição seja razoável e não discrimine com base em quem seriam os falantes ou pelo ponto de vista. Exemplos de fóruns não públicos incluem terminais de aeroportos, sistema de correio interno de uma escola pública e locais de votação.

É interessante observar que, no Brasil, o Supremo Tribunal Federal já teve a oportunidade de se manifestar sobre análises de fóruns pelo menos um par de vezes, sendo uma bastante recente no caso das manifestações de 2013.

Antes disso, ao julgar a ADI 1.969-4/DF, na qual era impugnado um decreto distrital editado a pretexto de regulamentar manifestações na Praça dos Três Poderes no Distrito Federal, o Relator Min. Ricardo Lewandowski enfrentou a matéria, construindo precedente paradigmático. Em suma, o

---

<sup>190</sup> Tradução livre do autor de trecho: "*The harder category, or subcategory, is the limited public forum. Doctrinally, this is where things start to get messy. Regardless of categorization, the limited public forum is a place or space designated for speech by "certain groups" or for "discussion of certain subjects." For example, a university can limit a public forum it establishes to use by student groups, and a school district can limit a public forum to the discussion of "school board business. (...) Moreover, the Supreme Court decisions in this area are almost always decided by 5-4 votes, making it even harder to predict how much leeway governments have to exert editorial control in a limited public forum.*" in Op. Cit. p. 4.

decreto buscava regulamentar o inciso XVI do art. 5º da Constituição, segundo o qual “[t]odos podem reunir-se pacificamente, sem armas, em locais abertos ao público, independentemente de autorização, desde que não frustrem outra reunião anteriormente convocada para o mesmo local, sendo apenas exigido prévio aviso à autoridade competente”, para - e aqui o conteúdo do decreto - proibir a utilização “de carros, aparelhos e objetos sonoros” na Praça dos Três Poderes e em outros espaços públicos elencados pela norma. O voto do Ministro Relator aplica análise do fórum de maneira bastante similar. Vejamos:

Ora, certo que uma manifestação sonora nas imediações de um hospital afetaria a tranquilidade necessária a esse tipo de ambiente, podendo, até mesmo, causar prejuízos irreparáveis aos pacientes. Ter-se-ia, nesse caso, uma hipótese de colisão entre direitos fundamentais, na qual o direito dos pacientes à recuperação da saúde certamente prevaleceria sobre o direito de reunião com tais características. Numa situação como essa, a restrição ao uso de carros, aparelhos e objetos sonoros mostra-se perfeitamente razoável. (...) a questão sob exame, no entanto, não guarda qualquer semelhança com tal hipótese. Na verdade, o Decreto distrital 20.098/99 simplesmente inviabiliza a liberdade de reunião e de manifestação, logo na Capital Federal, em especial na emblemática Praça dos Três Poderes, ‘*local aberto ao público*’, que, na concepção do genial arquiteto que a esboçou, constitui verdadeiro símbolo de liberdade e cidadania do povo brasileiro. Proibir a utilização ‘*de carros, aparelhos e objetos sonoros*’, nesse e em outros espaços públicos que o Decreto vergastado discrimina, inviabilizaria por completo a livre expressão do pensamento nas reuniões levadas a efeito nesses locais, porque as tornaria emudecidas, sem qualquer eficácia para os propósitos pretendidos. (...) Ademais, analisando-se a questão sob uma ótica pragmática, cumpre considerar que as reuniões devem ser, segundo a dicção constitucional, previamente comunicadas às autoridades competentes, que haverão de organizá-las de modo a não inviabilizar o fluxo de pessoas e veículos pelas vias públicas. Há que se ter em conta, por outro lado, que a utilização de aparelhos de som nas reuniões, que são limitadas no tempo, certamente não causará prejuízo irreparável àqueles que estão nas imediações da manifestação.

Em oportunidade mais recente, durante a profusão das manifestações do inverno de 2013 que tomaram as ruas do país, na Reclamação 15.887-MG o Ministro Luiz Fux também aplicou análise parecida justamente para garantir a aplicação do precedente firmado na ADI supracitada contra acordo do Tribunal de Justiça do Estado de Minas Gerais, que havia restringido

substancialmente o conteúdo do direito fundamental de livre manifestação do pensamento e de reunião.

Com efeito, a identidade material reside precisamente no fato de que tanto a decisão reclamada quanto o acórdão paradigma cuidam da constitucionalidade da proibição ao exercício do direito de reunião e de livre manifestação de pensamento em espaços públicos que, por suas características sociais e históricas, permitam a maior propagação das ideias e opiniões manifestadas pelos diversos segmentos da sociedade civil. Trata-se daquilo que o direito norte-americano intitulou como doutrina dos fóruns públicos (*public-forum doctrine*), segundo a qual uma sociedade livre deve criar uma plêiade de espaços nos quais se assegure, àqueles indivíduos que desejam se expressar, o direito de ter acesso aos lugares necessários para permitir a difusão da sua opinião entre as pessoas, notadamente aquelas áreas onde muitas delas se encontram. (...) Mas não é só. O direito de reunião consubstancia um componente indispensável à vida das pessoas e à própria existência de um substancial Estado Democrático de Direito. Conquanto a reunião de indivíduos em torno de determinados fins sociais tenha sempre existido no curso da história, é praticamente um consenso (...) que a vida associativa nos dias atuais encontra um solo fértil para as virtudes cívicas, ao mesmo tempo em que propicia uma base de sustentação para a construção de uma ordem democrática viável. (...) Trata-se, à evidência, de um direito moral, que deve ser reconhecido e protegido, independentemente de juízos morais meramente contingentes ou majoritários em uma determinada comunidade. Justamente por isso, sob um enfoque filosófico, a liberdade de reunião ostenta um status especial, um “peso absoluto”, com relação a razões de bem público, de cariz tipicamente utilitaristas, e a valores perfeccionistas, incompatíveis com o pluralismo existente nas sociedades contemporâneas. Com isso não se pretende afirmar que, sob o prisma jurídico-constitucional, o direito de reunião revista-se de caráter absoluto. Ao revés: o seu exercício pode encontrar-se limitado em virtude da colisão com o conteúdo de outros bens jurídicos de mesma estatura constitucional. Na realidade, o próprio constituinte originário previu expressamente uma restrição ao exercício do direito de reunião, quando decretado o Estado de Defesa (CRFB/88, art. 136, § 1º, I, alínea b).

É de suma importância frisar que, embora o direito de liberdade de expressão e a jurisprudência constitucional construída a partir dele possuam características bastante diferentes nos Estados Unidos e no Brasil, especialmente por cuidar de sistemas diferentes (Common e Civil Law), a lógica da análise do fórum também é feita pela Corte Constitucional brasileira, de modo que importar a mais longa experiência constitucional norte-americana não se mostra, de todo, uma comparação desmedida.

Portanto, voltando à decisão paradigmática<sup>191</sup> do caso Trump, a Corte Federal - não se tratando da Suprema Corte, é bom que fique claro - buscou considerar se um funcionário público pode, de acordo com a Primeira Emenda, “bloquear” uma pessoa de sua conta do Twitter em resposta às opiniões políticas que a pessoa expressou, e se a análise difere pelo fato de que, no caso, tratava-se do presidente do país. Para ambas as respostas, a Corte entendeu que não.

Em resumo, ao analisar o fórum em discussão, a Corte entendeu que partes da conta “@realDonaldTrump” - o “espaço interativo” em que os usuários do Twitter podem se envolver diretamente com o conteúdo dos tweets do Presidente - são devidamente analisadas sob as doutrinas do “fórum público” estabelecidas pela Suprema Corte norte-americana, segunda as quais tal espaço seria um fórum público designado, e que o bloqueio dos usuários com base em seu discurso político constitui uma discriminação de ponto de vista (*viewpoint discrimination*) que viola a Primeira Emenda. Vale a pena conferir o trecho em que a Corte determina por que a página pessoal do Twitter do Presidente se classifica como um fórum público designado:

Accordingly, we consider whether the interactive space is a designated public forum, with “governmental intent” serving as “the touchstone for determining whether a public forum has been created.” (...) “Intent is not merely a matter of stated purpose. Indeed, it must be inferred from a number of objective factors, including: [the government’s] policy and past practice, as well as the nature of the property and its compatibility with expressive activity.” (...) Here, these factors strongly support the conclusion that the interactive space is a designated public forum. “The @realDonaldTrump account is generally accessible to the public at large without regard to political affiliation or any other limiting criteria,” “any member of the public can view his tweets,” and “anyone [with a Twitter account] who wants to follow the account [on Twitter] can do so,” unless that person has been blocked. Similarly, anyone with a Twitter account who has not been blocked may participate in the interactive space by replying or retweeting the President’s tweets. Further, the account -- including all of its constituent components -- has been held out by Scavino as a means through which the President “communicates directly with you, the American people!” And finally, there can be no serious suggestion that the interactive space is incompatible with expressive activity: rather, Twitter as a platform is designed to allow users “to interact with other Twitter users in relation to [their tweets],” and users can use Twitter to “petition their elected representatives

---

<sup>191</sup> A decisão do caso está disponível no site da Suprema Corte norte-americana, mas pode ser facilmente acessada em: <https://knightcolumbia.org/sites/default/files/content/Cases/Wikimedia/2018.05.23%20Order%20on%20motions%20for%20summary%20judgment.pdf> (acessado em 04.02.19).



and otherwise engage with them in a direct manner,” Packingham, 137 S. Ct. at 1735. The interactivity of Twitter is one of its defining characteristics, and indeed, the interactive space of the President’s tweets accommodates a substantial body of expressive activity. Taking these factors together, we conclude that the interactive space of a tweet from the @realDonaldTrump account constitutes a designated public forum.

Considerando que o argumento utilizado para proibir o bloqueio foi o de que o “espaço interativo” da conta privada do Presidente constitui um fórum público, a decisão parece abrir caminho para a transposição de balizamentos regulatórios do mundo real para o virtual pela análise da função da conta dentro da rede social específica.

Neste ponto, é importante frisar que o Presidente utilizava a conta para divulgar conteúdo de interesse público acerca do seu governo. Portanto, ao entender que esse “espaço virtual” é um espaço público, podendo ser atingido pelas doutrinas do “fórum público”, segundo as quais restrições com base em conteúdo do discurso político são inconstitucionais sob o manto da Primeira Emenda, a decisão da Corte Federal pode representar um excelente embasamento tanto para eventuais regulações e julgamentos em casos concretos (pois dificilmente os parâmetros serão rígidos, devendo ser sopesados a cada hipótese), quanto para pensarmos um modelo de autorregulação regulada que empodere as intermediárias a governarem suas plataformas com base também nas restrições possíveis, a depender dos usuários e do contexto inserido.

## 5. A POSIÇÃO ALEMÃ: A NETWORK ENFORCEMENT ACT É UMA REGULAÇÃO BRANDA?

Em primeiro de outubro de 2017 entrou em vigor na Alemanha a *NetzDG*, uma lei<sup>192</sup> muito criticada por diversos setores da sociedade alemã e internacional, especialmente pela comunidade acadêmica preocupada com questões relativas aos direitos fundamentais, notadamente a liberdade de expressão. A lei pode ser caracterizada como uma norma com quatro objetivos principais, os quais serão brevemente expostos a seguir para que se entenda a ideia geral da lei. Em seguida, comenta-se sobre as principais objeções feitas à norma, para depois traçarmos um panorama geral do *performance standard* e do *compliance regime* desenhados pela legislação.

O primeiro objetivo da lei é melhorar o *enforcement* de legislações do mundo físico no mundo digital. Por isso mesmo a lei se chama “*Das Netzwerkdurchsetzungsgesetz*” (*Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken – NetzDG*), ou, em inglês “*Act to Improve Enforcement of the Law in Social Networks*” (*Network Enforcement Act*). Percebe-se, então, que a lei está preocupada em melhorar a aplicação de legislação já existente no mundo digital. Nesse sentido, há um nítido foco para as práticas de ilícitos que preenchem determinados tipos penais do Código Alemão de Direito Penal.<sup>193</sup> Portanto, a lei visa combater *harms* já tipificados.

O segundo objetivo limita-se a obrigar as empresas de tecnologias - que operam as maiores plataformas interativas - a prestarem informações relativas aos ilícitos praticados na rede, as medidas tomadas para impedi-los, e a quantidade e o escopo das reclamações enviadas às empresas pelos usuários. Com isso, a lei busca aumentar a transparência dessas empresas e melhor informar a sociedade sobre o que acontece nestas “*caixas pretas*”,

---

<sup>192</sup> A lei está disponível em língua inglesa em:

<[https://www.bmju.de/SharedDocs/Gesetzgebungsverfahren/Dokumente/NetzDG\\_engl.pdf?\\_blob=publicationFile&v=2](https://www.bmju.de/SharedDocs/Gesetzgebungsverfahren/Dokumente/NetzDG_engl.pdf?_blob=publicationFile&v=2)> (acessado em 04.02.19).

<sup>193</sup> Art. 1 §1º (3) Unlawful content shall be content within the meaning of subsection (1) which fulfils the requirements of the offences described in sections 86, 86a, 89a, 91, 100a, 111, 126, 129 to 129b, 130, 131, 140, 166, 184b in connection with 184d, 185 to 187, 241 or 269 of the Criminal Code and which is not justified.

mantendo-se, ao mesmo tempo, a tradição de incentivo ao desenvolvimento tecnológico, porque aplicada apenas às maiores plataformas.

O terceiro objetivo consiste em regulamentar os processos internos das empresas no que tange ao tratamento de reclamações sobre a prática desses ilícitos nas redes sociais, tentando criar uma transparência e proceduralização do direito dentro desses procedimentos internos. Este talvez seja o objetivo mais complexo e criticado, porque contém a obrigação de retirar ou bloquear conteúdo evidente e manifestamente ilícito do ar em 24 horas após o recebimento da reclamação, e fazer o mesmo para os demais conteúdos ilícitos (mais difíceis de serem avaliados - “zona cinzenta”) em prazo mais alargado, de 7 dias. De certa forma, a principal crítica veiculada aqui está ligada à aplicação do Direito por entes privados.

O quarto objetivo principal da lei é estabelecer sanções administrativas e pecuniárias às empresas que descumprirem os comandos legais. As multas podem variar de 5 milhões para pessoas, a 50 milhões de euros para empresas. Neste ponto, as principais críticas feitas são as de que a lei parece depositar nas plataformas a responsabilidade de determinar se um conteúdo específico viola ou não a lei, o que se mostra uma situação problemática, pois a análise de discursos, especialmente de conteúdo político, exige uma compreensão aguçada do contexto cultural e legal, sendo certamente difícil de ser feita em 24 horas, ou em 7 dias para casos menos evidentes. Assim, a existência de multas milionárias pode acabar forçando as empresas a, na dúvida, optarem por retirar do ar. Isto, segundo os críticos, levaria ao *overblocking*.

Antes de adentrarmos em algumas questões interessantes da lei, ressalta-se que embora tenha eficácia apenas em solo alemão, a lei pode servir como um teste para outras democracias ocidentais, especialmente o caso brasileiro, pois nosso modelo de responsabilização civil dos provedores de aplicação é similar ao alemão, e agora a Lei Geral de Proteção de Dados também se assemelha muito - tendo sido inclusive inspirada - ao GDPR (*General Data Protection Regulation*). Como dito, a lei vem sendo alvo de severas críticas na sociedade, e segundo o Human Rights Watch, estabelece

um precedente perigoso para governos que buscam restringir o discurso online, forçando as empresas a censurar em nome do Estado.<sup>194</sup>

Antes de passar aos pontos-chave da lei, cumpre explicar as objeções centrais, visando a uma análise posterior objetiva.

---

<sup>194</sup> Sobre o tema ver: <<https://www.hrw.org/news/2018/02/14/germany-flawed-social-media-law>> (acessado em 04.02.19).

## 5.1 Principais objeções à NetzDG: *overblocking* e aplicação privada do direito

Pretende-se neste tópico analisar conjuntamente alguns recentes artigos sobre a *NetzDG*, dando-se preferência aos autores alemães, sem, é claro, abrir mão de algumas considerações e comparações ao ordenamento brasileiro. As principais obras consultadas são: “A lei alemã para a melhoria da aplicação da lei nas redes sociais”,<sup>195</sup> de Martin Eifert, professor catedrático de direito público e administrativo da Humboldt-Universität; “Regulating Intermediaries to Protect Privacy Online – the Case of the German NetzDG”,<sup>196</sup> de Wolfgang Schulz, professor de Direito Público, Mídia e Teoria do Direito da Universidade de Hamburgo; e “A mudança na esfera pública pela inteligência artificial” de Thomas Vesting, catedrático da Goethe-Universität.

Em suma, as principais preocupações quanto à lei nos levam ao que Martin Eifert chamou de “*pontos materiais centrais críticos*” da lei, o perigo de *overblocking* e a aplicação privada do direito.<sup>197</sup>

Segundo Eifert, as críticas de *overblocking* não se sustentam levando-se em conta a estrutura da lei, porque ela estabelecerá (i) *standards* materiais para a avaliação de conteúdos, e (ii) uma estrutura de incentivos fáticos e ameaças de punição apenas por erros sistêmicos; e não para casos pontuais de erros de bloqueio. Por questões de objetividade, a passo que se expõe as principais objeções feitas à lei, comenta-se também as respostas dadas às críticas.

A primeira crítica é que a obrigação contida no Art. 1º § 3º (1) 2 da *NetzDG*, compele as plataformas a bloquear ou remover conteúdo manifesta e

---

<sup>195</sup> EIFERT, Martin. *A Lei Alemã para a melhoria da aplicação da lei nas redes sociais (NetzDG) e a regulação da plataforma*, op.cit. nota 104.

<sup>196</sup> SCHULZ, Wolfgang, *Regulating Intermediaries to Protect Privacy Online – The Case of the German NetzDG* op. cit. nota 133.

<sup>197</sup> Na opinião do pesquisador, as objeções sobre a lei poder gerar *overblocking* são discutíveis, contudo, não se sustentariam quando considerada a abordagem base da lei. Já em relação à problemática da aplicação privada do direito, segundo ele, o ponto de crítica estaria fundamentalmente incorreto. In EIFERT, Martin. *A Lei Alemã para a melhoria da aplicação da lei nas redes sociais (NetzDG) e a regulação da plataforma*, op.cit. nota 104 . p. 73.

evidentemente ilícito (“obviamente ilegal”) dentro de 24 horas, o que teoricamente criaria uma tendência na empresa de tecnologia a, por reduções de custo, bloquear. Por esse motivo, os defensores da liberdade de expressão argumentam que, em primeiro lugar, haveria dúvidas sobre se, de fato, há tanta obviedade assim ao se decidir se um conteúdo é manifestamente ilegal ou não, dado que o contexto cultural e temporal devem sempre ser levados em conta, especialmente porque os grandes temas (*trend topics*) geram milhões de conteúdos por dia. Segundo, cada conteúdo sinalizado deveria ser avaliado para identificar as partes “obviamente ilegais” e as que não seriam, mas o que acontece é o bloqueio do conteúdo por completo. Portanto, considerando que, de acordo com o Facebook, somente a difamação e o discurso de ódio são responsáveis por mais de 100 mil bloqueios por mês na Alemanha, dada essa quantidade de bloqueios e a dificuldade de gerir cada caso individualmente, a tendência seria que um provedor derrubasse, em caso de dúvida, qualquer conteúdo sinalizado, apenas para economizar custos.

Além disso, o prazo de sete dias para conteúdo “na zona cinzenta” (não obviamente ilegal) também causaria transtornos. Como explica Wolfgang Schulz:

A avaliação quanto a uma fala ser uma declaração de fato ou uma proclamação de uma opinião é essencial para uma avaliação sob a lei alemã. Essa é uma questão complexa, e pode ser até que tribunais diferentes discordem sobre o resultado. O mesmo é verdadeiro para a questão de saber se uma declaração de fato é evidentemente verdadeira ou não. Conduzir tais avaliações dentro do prazo determinado coloca pressão sobre um provedor e pode empurrá-lo novamente para a solução simples, porém, adversa aos direitos humanos, para derrubar o conteúdo em quase todos os casos. Além disso, os provedores não possuem as informações sobre o contexto - e as ferramentas de coleta de informações necessárias - para fazer uma avaliação adequada.<sup>198</sup>

---

<sup>198</sup> Tradução livre do autor do trecho: “*The assessment whether a speech is a statement of fact or a proclamation of an opinion is essential for an assessment under German law. This is a complex issue, and it might be that even different courts disagree on the result*27. *The same is true for the question whether a statement of fact is evidentially true or not. To conduct such assessments within the given time frame puts pressure on a provider and might again push it to the simple but human rights-adverse solution to take down the content in almost any case. Furthermore, the providers lack the information about the context – and the necessary information gathering tools – to make a proper assessment.*” In <sup>198</sup> SCHULZ, Wolfgang, *Regulating Intermediaries to Protect Privacy Online – The Case of the German NetzDG* op. cit. nota 133. p. 8.

Como reporta Schulz, a *NetzDG* não ultrapassaria o teste da proporcionalidade, porque, embora reduzir crimes na Internet seja um objetivo legítimo, as disposições da lei não ajudariam a identificar o culpado, e mesmo que se considere apenas a redução dos crimes, a justificativa formal para a criação da lei foi a necessidade de se manter uma cultura de debate político e não a diminuição de crimes.

Para Schulz, “o desejo de proteger a cultura política - plausível como é - não é suficiente para limitar os direitos humanos”.<sup>199</sup> Outro ponto de crítica é que os órgãos estatais aos quais as plataformas precisam se reportar - e que são os órgãos autorizados a aplicar sanções - seriam diretamente vinculados ao Ministério da Justiça alemão, e isso seria especialmente problemático, uma vez que a independência da imprensa e da mídia na Alemanha são altamente protegidas pelo Tribunal Constitucional Federal alemão, remontando às razões do uso das mídias durante a ditadura nazista.<sup>200</sup>

Importante ter em vista que a história da Alemanha torna todo e qualquer debate sobre discursos políticos, especialmente o discurso de ódio, um tanto quanto sofisticado. Portanto, considerando que a lei foi criada dentro de um contexto de crescimento da extrema direita na Alemanha (e na Europa como um todo), e considerando ainda as notícias de tentativa de outros países, como a Rússia, de interferirem na opinião pública dos países europeus, a lei também deve ser enxergada como uma resposta normativa rápida e responsiva a esses problemas.

Também por esses motivos, em defesa da lei, Martin Eifert formula argumentos bastante sólidos, no sentido de que os *standards* materiais para a avaliação de conteúdos ilícitos estipulados pela lei e pelos termos de uso das plataformas, bem como a estrutura de incentivos fáticos de mercado e as ameaças punitivas apenas para hipóteses de falhas sistêmicas, acabam desempenhando um papel crucial na validade da lei.

Em primeiro lugar, Eifert argumenta que os termos de uso (*community standards*) das plataformas geralmente são o critério mais utilizado para bloqueio de conteúdo, e eles são claramente mais restritivos à liberdade de

---

<sup>199</sup> Idem, *Ibidem*. p. 9.

<sup>200</sup> Lembra-se aqui a discussão já travada sobre a independência da Autoridade Nacional de Proteção de Dados criada no apagar das luzes do Governo Temer.

expressão do que a própria lei, como, por exemplo, em assuntos relacionados ao decoro e moral sexual (nudez, etc). Os termos de uso são visivelmente menos permissivos do que a lei. Então, as discussões sobre *overblocking* e censura não estariam conectadas necessariamente apenas à lei, mas ao sistema que se utiliza de termos de uso para bloquear como um todo. Além disso, em comparação com proibições jurídicas baseadas em fundamentações históricas incidentes na Alemanha (negação do holocausto, por exemplo), os termos de uso das plataformas costumam privilegiar a liberdade de expressão, aproximando-se bastante da cultura norte-americana do *free speech*, segundo a qual os *standards* devem ser o mais amplos possíveis para evitar a particularização e a formação de guetos de comunicação na Internet.

Para completar, Eifert lembra que o direito de autodeterminação das plataformas, ou seja, o fato de elas poderem estabelecer suas próprias regras, também é formado por direitos fundamentais que igualmente devem ser ponderados com a liberdade de expressão. Nesse sentido, a discussão iria invariavelmente ser direcionada ao controle judicial dos termos de uso. Eifert lembra que, segundo o Código Civil Alemão (§ 307 I 1 BGB), em uma análise estritamente axiológica em caso de ponderação, os contratos de serviço das redes sociais provavelmente seriam inválidos, pois o usuário estaria desproporcionalmente desfavorecido em relação à empresa, sob vários aspectos. Algo bastante parecido com a ideia dos contratos de adesão e a lógica protetiva do Código de Defesa do Consumidor no direito brasileiro.

Em segundo lugar, Eifert utiliza um argumento que vem ao encontro dos temas debatidos neste trabalho, qual seja: a mesma abordagem daqueles que enxergam as redes sociais como fóruns públicos. Tal conceptualização permitiria enfatizar outros valores constitucionais em relação ao uso dos espaços (interesses públicos x privados). A conclusão é que quaisquer abordagens resultam em ponderações de interesses conflitantes, permanecendo ainda em aberto a pergunta sobre quais seriam os resultados concretos.

Por fim, o pesquisador coloca dois argumentos de ordem mercadológica, de maneira que seria necessário averiguar se a postura adotada pela *NetzDG* necessariamente cria um perigo sistêmico de bloqueios de conteúdo. A uma, Eifert ressalta que os intermediários não se encontram na



posição do comunicador nem do ouvinte, possuindo um interesse econômico que envolve manter o maior número possível de usuários como fornecedores de dados e recebedores de publicidade. Além disso, as plataformas precisam lidar com a necessidade de manter uma imagem positiva da empresa para terem aceitação no mercado. Portanto, se as plataformas agirem conservadoramente na contenção de discursos de ódio, por exemplo, elas serão acusadas de serem coniventes com a fala, o que afastará alguns usuários. Por outro lado, caso retirem conteúdos demais, serão acusadas de repressoras e enviesadas, o que afastará outros. Ambas as acusações seriam danosas para a empresa, de modo que valeria a pena fazer um bom trabalho e manter-se equidistante.

A duas, Eifert argumenta que a *NetzDG* estipula como *performance standard* e prazos para o apagamento vinculados apenas a conteúdos obviamente ilícitos, e as multas estão mais vinculadas à necessidade de se instaurar um procedimento efetivo e transparente, do que ao bloqueio de conteúdos específicos em si. Além disso, a multa pecuniária só incide quando há um *overblocking* sistêmico, de forma que a incidência de sanções pecuniárias não levariam as plataformas a terem esse ou aquele viés, mas sim a realizarem um balanceamento de interesses, pois as multas estão mais afetadas ao descumprimento de procedimentos, ou quando eles não são observados corretamente, do que por um ou outro bloqueio equivocado. Para concluir seu argumento, o professor utiliza o exemplo das obrigações de bloqueio de conteúdo de Direitos Autorais nos Estados Unidos, as quais demonstram, há quase 20 anos, que os problemas de *overblocking* estão relacionados aos casos em que há uso de procedimentos automatizados. Para ele, o modelo de “*notice and take down*” dos Direitos Autorais é visto amplamente como um ponto central e paradigmático da aplicação do direito.

Em relação à segunda objeção posta, qual seja, a de que obrigar as plataformas a distinguirem conteúdo lícito do ilícito equivaleria a permitir uma aplicação privada do Direito, o problema arguido é que a aplicação do Direito é monopólio estatal e, portanto, não pode ser delegado às empresas de tecnologia. Assim, a lei seria inconstitucional porque violaria os fundamentos do Estado de Direito.

Para Eifert, as críticas formuladas nesse sentido seriam muito barulho por nada. A primeira razão colocada tem a ver com o fato de o apagamento ou bloqueio de conteúdo ilícito ser uma consequência jurídica normal resultante da rede de relações travadas nas redes sociais que possuem caráter privado. Nesse sentido, Eifert argumenta que, embora a lei crie uma “responsabilidade civil por monitoramento” (*Störerhaftung*), ou seja, como já explicamos acima, a possibilidade de a empresa ser responsabilizada caso viole o *duty to care* - a obrigação de ser suficientemente diligente -, a jurisprudência do Tribunal Federal Constitucional “*basicamente cria obrigações no que toca à causação ou manutenção de estados de coisas ilícitas*”, o que não penaliza o provedor por condutas isoladas.

A lógica seria mais ou menos a mesma caso comparássemos as redes sociais com os meios clássicos de comunicação: o ofendido se voltaria contra o ofensor, normalmente pelas vias judiciais, e o meio de comunicação estaria no meio dessa relação. Os intermediários, considerados pontos de gargalo da comunicação, apresentam-se em posição favorável para regulações jurídicas, porque a jurisprudência sobre a responsabilidade indireta de monitoramento (*breach of duty to care*) - muito embora apresente variações, pois foi desenvolvida em distintos campos do direito, como a concorrência, direitos da personalidade e direitos autorais -, exige não apenas uma mera omissão, mas a comprovação do não cuidado, a causação ou a manutenção da “perturbação”. Portanto, desde que a empresa demonstre que foi suficientemente diligente, ela não será responsabilizada.

Vejamos, por exemplo, no caso dos direitos autorais. Na hipótese de transmissão de conteúdo protegido, os intermediários são geralmente apenas auxiliares do autor da violação de direitos autorais e não são responsáveis por uma infração direta. Mesmo se os intermediários forem os causadores indiretos ou contribuintes da violação de direitos autorais, seu dever é limitado às obrigações de fazer cessar e remover, de modo que, para que os intermediários sejam responsabilizados, eles precisariam não ter sido suficientemente diligentes. Em suma: (i) o intermediário deve, de alguma forma, contribuir voluntária e causalmente para a violação de um interesse legal protegido. Não depende de culpabilidade (responsabilidade objetiva); (ii) além disso, o intermediário deve ter a possibilidade legal de ter impedido a ofensa

praticada. Normalmente, isso ocorre porque eles podem impedir o ato ilícito simplesmente deixando de disponibilizar seu serviço ao infrator.<sup>201</sup>

No entanto, a fim de evitar que o violação do *duty to care* se tornasse excessiva, a jurisprudência alemã introduziu um terceiro requisito como corretivo: a violação das obrigações de *due diligence* exige um equilíbrio global de interesses e uma alocação de risco de julgamento quanto a se o cumprimento da diligência devida era razoável. Ou seja, a Corte geralmente requer que a violação de direitos autorais tenha sido aparente. Isso quer dizer, deve ser uma infração facilmente reconhecível ou a violação deve ter sido informada pela parte violada. Essa parece ser a mesma lógica pretendida pela *NetzDG* ao exigir a retirada em 24 horas de conteúdo “*obviamente ilegal*”. Agora, isso pode significar não só uma obrigação de devida diligência para impedir a violação específica, mas também uma obrigação de diligência para evitar outras violações óbvias do mesmo tipo.<sup>202</sup>

Por fim, uma última ponderação feita por Eifert, para contrariar aqueles que dizem que a *NetzDG* permite uma aplicação privada do direito, é que a persecução penal estatal dos autores de conteúdos ilegais permanece intocada, porque a lei prevê expressamente a guarda dos conteúdos para fins de prova por 10 semanas, indicando que a persecução jurídica não se limita à retirada de conteúdos.

---

<sup>201</sup> CONRAD, Albrecht. *Exceptions to copyright protection and the permitted uses of copyright works in the hi-tech and digital sectors*. Disponível em: <[https://www.aippi.org/download/committees/216/GR216germany\\_en.pdf](https://www.aippi.org/download/committees/216/GR216germany_en.pdf)>(acessado em 04.02.19).

<sup>202</sup> Idem, Ibidem. p. 6. Em trecho interessante sobre a jurisprudência do Tribunal Federal Constitucional: “*Furthermore, on the question as to when compliance with a due diligence obligation can be reasonably expected of the party subject to a claim for breach of duty of care, the German Federal Court of Justice has considerably eased the right holder's burden of proof. In principle, the burden of proof fundamentally lies with the claimant (and thus the right holder). However, there is a secondary burden of proof and stating the case for the party claimed against, as this party must usually be in possession of relevant knowledge of its technical infrastructure. The party subject to the claim is thus obligated to argue in detail which protection measures it is able to take and why it is unreasonable to expect them of it (BGH GRUR 2008, 1097 para. 19 et seq. - Namensklau im Internet, as Fromm / Nordemann /Jan Bernd Nordemann cited above, § 97 para. 162 at the end).*”

## 5.2 A configuração legislativa da *NetzDG* e suas principais características

No art. 1º §1º seções (1), (2) e (3) da *NetzDG* temos o âmbito de aplicação da norma. A *NetzDG* se aplica a todos “os provedores de serviços de telecomunicação que, com cunho lucrativo, operam plataformas na internet que funcionam de modo que usuários possam compartilhar conteúdos indeterminados com outros usuários ou torná-los acessíveis ao público (redes sociais). Plataformas de cunho jornalístico e redacional que sejam de responsabilidade dos provedores de serviço não são consideradas como redes sociais no sentido desta lei. O mesmo é válido para plataformas que são direcionadas para comunicação individual ou para a difusão de conteúdos específicos”. Além disso, “o provedor de uma rede social está liberado das obrigações [...] quando a rede social tiver, domesticamente, menos do que dois milhões de usuários registrados.”<sup>203</sup>

Quanto ao âmbito de aplicação, a lei é nitidamente direcionada às empresas de tecnologia que operam as grandes redes sociais como o Facebook, o Twitter e o Youtube, mantendo a tradição alemã de incentivo ao desenvolvimento tecnológico. Inclusive, a lei parece resolver a discussão travada em momento anterior, acerca do fato da transposição da mídia tradicional para os veículos digitais ter feito com que muitas dessas entidades oferecessem serviços híbridos e, por isso, acabarem se encaixando como novos intermediários de uma forma ou de outra. A lei aparentemente resolve a discussão sobre os intermediários híbridos ao excluir do âmbito de validade as plataformas de cunho jornalístico e redacional ou plataformas direcionadas a conteúdos específicos.

---

<sup>203</sup> A tradução da *NetzDG* utilizada aqui consta na obra ABBOUD, Georges. NERY Jr., Nelson. e CAMPOS, Ricardo (Coordenadores). *Fake News e Regulação*. Coleção Direito do Estado e Transformação. São Paulo: Thomson Reuters Brasil, 2018. p 233-241.

Quanto aos conteúdos ilícitos, a lei arrola expressamente os tipos do Código Penal alemão que constituem os danos a que se busca evitar.<sup>204</sup> Em geral, *“as infrações penais se referem a conteúdos de comunicação que (...) violam direitos da personalidade de terceiros ou perturbam a paz pública pela degradação de grupos determinados.”*<sup>205</sup>

Para melhor compreender quais tipos de crimes estão no âmbito da lei e aprofundar o debate sobre a mesma, vale a pena consultar o que aconteceu na prática, logo nos primeiros semestres de vigência da lei. Vejamos os crimes dispostos no primeiro relatório entregue pelo Facebook e o número de denúncias feitas:

---

<sup>204</sup> Os tipos penais são: Seção 86 – disseminação de material de propaganda de organizações inconstitucionais; Seção 86a – utilização de símbolos de organizações inconstitucionais; Seção 89a – preparação de crime grave contra o estado; Seção 91 – incitação à prática de crime grave contra o estado; Seção 100a – falsificação; Seção 111 – incitação pública à prática de crime; Seção 126 – perturbação da paz pública com ameaças de cometimento de crimes; Seção 129 a 129b – formação de organizações terroristas ou criminosas; Seção 130 – incitação ao ódio; Seção 131 – disseminação de imagens de violência; Seção 140 – premiação ou aprovação de atos criminosos; Seção 166 – difamação de religiões e de associações ideológicas ou religiosas; Seção 184b em conjunto com 184d – distribuição, aquisição e posse de pornografia infantil; Seção 185 a 187 – insulto e difamação; Seção 201a – violação da intimidade em consequência de fotografias; Seção 241 – ameaça à prática de crime; Seção 269 – falsificação de dados com o intuito de fornecer prova. Informação disponível em: <<https://www.facebook.com/help/285230728652028>> (acessado em 04.02.19).

<sup>205</sup> EIFERT, Martin. *A Lei Alemã para a melhoria da aplicação da lei nas redes sociais (NetzDG) e a regulação da plataforma*, op.cit. nota 104. p 66.

Criminal Code Provision	Reporter Type		
	Reports from Complaint Bodies	Reports from Other Individuals	Total
Dissemination of propaganda material of unconstitutional organizations (§ 86)	11	53	64
Using symbols of unconstitutional organizations (§ 86a)	13	53	66
Preparation of a serious violent offense endangering the state (§ 89a)	4	20	24
Encouraging the commission of a serious violent offense endangering the state (§ 91)	2	24	26
Treasonous forgery (§ 100a)	5	22	27
Public incitement to crime (§ 111)	14	109	123
Breach of the public peace by threatening to commit offenses (§ 126)	9	74	83
Forming criminal or terrorist organizations (§§ 129 - 129b)	4	26	30
Incitement to hatred (§ 130)	37	210	247
Dissemination of depictions of violence (§ 131)	17	69	86
Rewarding and approving of offenses (§ 140)	12	83	95
Defamation of religions, religious and ideological associations (§ 166)	13	79	92
Distribution, acquisition, and possession of child pornography (§ 184b in conjunction with § 184d)	3	16	19
Insult (§ 185)	56	404	460
Defamation (§ 186)	46	361	407
Intentional defamation (§ 187)	38	304	342
Violation of intimate privacy by taking photographs (§ 201a)	17	78	95
Threatening the commission of a felony (§ 241)	16	103	119
Forgery of data intended to provide proof (§ 269)	14	42	56

Entre 1º de janeiro e 30 de junho de 2018, o Facebook recebeu 886 reclamações baseadas nas provisões da *NetzDG*. O relatório produzido<sup>206</sup> pela empresa tem apenas 8 páginas, é simples e objetivo, e parece mostrar que a lei alemã foi cumprida. Ao consultá-lo, a percepção é de que a empresa foi capaz de criar um sofisticado procedimento interno que responde objetivamente às obrigações trazidas pela lei, ao mesmo tempo que protege a empresa de responsabilizações, justamente pelo cumprimento do dever de diligência. Nesse sentido, dentro do procedimento de reclamação há a constituição de material probatório, depositando-se no reclamante boa parte do ônus de produzir as provas do material ilícito e as razões do por quê haveria a

<sup>206</sup> De fato, ao que parece, as críticas quanto à impossibilidade de tratar todas as reclamações não se sustentam. 886 reclamações em 6 meses de validade da lei não se mostram uma quantidade excessiva. Talvez porque o próprio procedimento para “denúncia” criado pela empresa é bastante sofisticado. O relatório entregue pelo Facebook está disponível em: <[https://fbnewsroomus.files.wordpress.com/2018/07/facebook\\_netzdg\\_july\\_2018\\_english-1.pdf](https://fbnewsroomus.files.wordpress.com/2018/07/facebook_netzdg_july_2018_english-1.pdf)> (acessado em 04.02.19).

subsunção da conduta no tipo penal. O mecanismo pelo qual os usuários podem relatar o conteúdo mostra-se, em certa medida, bastante complexo.<sup>207</sup>

Já os relatórios entregues pelo Youtube,<sup>208</sup> rede social de compartilhamento de vídeos, tiveram que lidar com um número consideravelmente maior de “denúncias”. Por exemplo, entre julho e dezembro de 2018 - portanto, o segundo relatório entregue pela empresa - houve cerca de 250 mil itens denunciados. No entanto, o Youtube considera como “itens” tanto vídeos postados como comentários. Dessas 250 mil denúncias, cerca de 55 mil resultaram em ações da empresa: ou o conteúdo violava as diretrizes da comunidade (*community standards*) e foi removido globalmente ou o conteúdo foi identificado como ilegal de acordo com um dos 22 estatutos do código penal alemão relacionado à *NetzDG*, e foi apenas restringido localmente (*geo blocked*).

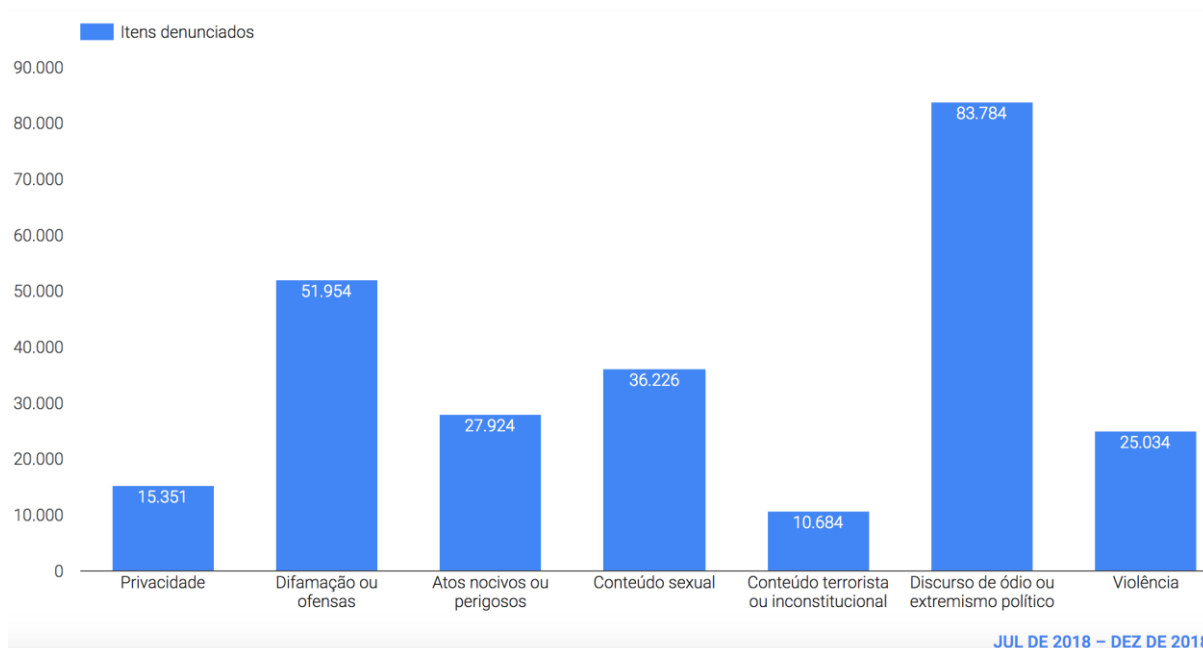
O Google, empresa que controla o Youtube, destaca que 78,23% dos casos denunciados não levaram a nenhuma medida, porque o conteúdo não violava as diretrizes da comunidade nem aos estatutos da *NetzDG*, sendo que em 95,22% dos casos, uma medida foi tomada em até 24 horas do recebimento. Na maioria desses casos, o conteúdo é removido globalmente devido a violações das diretrizes da comunidade do YouTube e não por causa da *NetzDG*. O Youtube faz uma interessante divisão em categorias de *harms*:

---

<sup>207</sup> O Facebook pede que os “denunciadores” forneçam as seguintes informações: (i) informações de contato completas, incluindo se o ofensor é uma autoridade oficial (*Beschwerdestelle* na *NetzDG*), (ii) links para o conteúdo específico no Facebook, se disponível, (iii) Seção(s) do Código Penal alemão supostamente violada pelo conteúdo relatado, (iv) declarações ou imagens específicas no conteúdo denunciado que são supostamente ilegais na *NetzDG*, (v) por que essas declarações ou imagens são consideradas ilegais na *NetzDG*, (vi) uma ordem judicial, se disponível, (vii) uma assinatura eletrônica.

<sup>208</sup> Relatórios disponíveis em:

<[https://transparencyreport.google.com/netzdg/youtube?items\\_by\\_submitter=period:1&lu=items\\_by\\_submitter](https://transparencyreport.google.com/netzdg/youtube?items_by_submitter=period:1&lu=items_by_submitter)> (acessado em 04.02.19).



Em relação às obrigações de prestar informações contidas no art. 1º § 2º da *NetzDG*, as redes sociais devem apresentar os supracitados relatórios a cada 6 meses, devendo ser de fácil reconhecimento, e acessível de forma direta e contínua.

O relatório deve abordar no mínimo os seguintes aspectos: (i) considerações gerais sobre quais esforços o provedor tomou para impedir ações ilícitas na plataforma; (ii) apresentação dos mecanismos de denúncia dos conteúdos ilícitos além dos critérios de decisão; (iii) número de reclamações relativas a conteúdos ilícitos, dividindo-os em reclamações provenientes de usuários ou de gabinetes oficiais, categorizando-as segundo o motivo da denúncia; (iv) a organização, recursos de pessoal, competência técnica e linguística das unidades responsáveis pela análise das denúncias, (v) associação ou filiação a associações profissionais ou comerciais das quais a empresa faz parte,<sup>209</sup> indicando se existe um setor de reclamações nessas associações; (vi) número de reclamações nas quais uma agência ou entidade externa foi consultada para se chegar a uma decisão; (vii) número de reclamações recebidas que levaram ao apagamento, retirada ou bloqueio de conteúdo; (viii) o tempo decorrido entre a reclamação e a ação adotada pela empresa; (ix) as medidas tomadas para orientação e instrução sobre a decisão.

<sup>209</sup> Essa obrigação fica pouco clara no texto da lei, mas recorrendo aos relatórios das empresas, percebe-se que a vontade do legislador foi saber a que tipo de associações profissionais e comerciais a empresa está conectada, que tipo de parcerias com entidades a empresa faz e como cada uma delas opera suas reclamações.



Acerca do tratamento de reclamações sobre conteúdos ilícitos, o art. 1º §3º da *NetzDG* cria determinações organizacionais, obrigações de prestar informações e prazos. Como explica Martin Eifert, ao exigir apenas o estabelecimento de um procedimento de comunicação e envio de reclamações sobre conteúdos ilícitos que seja de fácil reconhecimento, diretamente acessível e constantemente disponível para o usuário, a lei renuncia a criar, ela própria, um *compliance-regime* específico. Ou seja, com os *performance-standards* criados, a lei não faz uma configuração pormenorizada de como deve ser o procedimento de cada empresa, assegurando altos graus de liberdade para as plataformas em sua implementação. Nessa toada, basta comparar os relatórios supramencionados (Facebook e Youtube) com o relatório do Twitter,<sup>210</sup> e reparar como cada plataforma teve a liberdade para criar seu próprio procedimento, que deve ser monitorado e supervisionado pela direção da plataforma por meio de controles mensais. Além disso, a lei também prevê que a empresa envie a decisão sobre remoção a uma instituição (*Zustellungsbevollmächtigten*) de regulação da autorregulação. Se trata do destinatário autorizado para o recebimento de processos civis e administrativos.

A tabela abaixo, retirada de artigo de pesquisadores do Alexander von Humboldt Institut für Internet und Gesellschaft (HIIG) referenciado em notas, fornece uma visão geral do número total de itens mencionados nos relatórios de transparência das principais redes sociais durante os seis primeiros meses de validade da lei, bem como a porcentagem de casos em que as plataformas cumpriram as reclamações e, subsequentemente, removeram ou bloquearam o conteúdo.

---

<sup>210</sup> O relatório do Twitter está disponível em: <<https://cdn.cms-twdigitalassets.com/content/dam/transparency-twitter/data/download-netzdg-report/current-report.pdf>> (acessado em 04.02.19)).

**Table 1: Reported numbers by selected platforms, January – June 2018**

Platform	Total items reported	Reports resulted in action (removal rate)	Removal rate within 24h
Facebook	1,704	362 (21.2 percent)	76.4 percent (of reports)
YouTube	241,827	58,297 (27.1 percent)	93.0 percent (54.199)
Google+	2,769	1,277 (46.1 percent)	93.8 percent (1.198)
Twitter	264,818	28,645 (10.8 percent)	97.9 percent (28.044)

Agora, o que esses números nos relatórios realmente nos contam?

Embora esses números agora estejam públicos para debate e também possam levar a novas pesquisas, é importante ponderar que qualquer análise com base em relatórios estatísticos, índices, classificações, e interpretações de números nus podem produzir falhas.

Em primeiro lugar, os números nus são difíceis de serem verificados porque os pesquisadores não têm acesso aos "dados brutos". Muitas vezes, relatórios com números crus podem ocultar mais do que revelar, pois eliminam informações importantes sobre o contexto. Em segundo lugar, de acordo com os pesquisadores do HIIG,<sup>211</sup> os relatórios atuais não fornecem uma medida valiosa para comparar o desempenho entre as plataformas, pois cada um dos relatórios segue sua própria estrutura de organização - conforme liberdade permitida pela lei.

Por outro lado, em defesa das plataformas, os relatórios parecem cumprir exatamente o que determina a lei e provavelmente boa parte do procedimento de avaliação utiliza uma combinação de máquinas e inteligência artificial com trabalho humano,<sup>212</sup> de sorte que o próprio processo constitui uma tecnologia, e pode estar protegida por leis de propriedade industrial. Assim, não faz muito sentido que a empresa os “entregue de graça”.

<sup>211</sup> GOLLATZ, Kirsten. RIEDL, Martin J. POHLMANN, Jens. *Removals of online hate speech in numbers*. Disponível em: <<https://www.hiig.de/en/removals-of-online-hate-speech-numbers/>> (acessado em 04.02.19).

<sup>212</sup> No relatório do YouTube há uma explicação pormenorizada de como funcionam os procedimentos com humanos e máquinas.

Em suma, a questão discutida pelos pesquisadores é a de que, caso as plataformas continuem desenvolvendo relatórios muito distintos e ocultem informações sobre os “dados brutos”, não será possível comparar o desempenho entre elas. Ademais, critica-se também o fato de as empresas passarem a moldar o conceito de “transparência” - que será sempre o mínimo possível, via o “*friendly talk*” comum das empresas de tecnologia. Se não houver práticas de relatórios padronizadas e os usuários tiverem que passar por diferentes obstáculos em cada plataforma para registrar uma denúncia, será que esses números realmente revelam alguma coisa?

Por fim, dentro das nossas pretensões de traçar um panorama geral da *NetzDG* objetivamente, o último ponto crucial a ser comentado tem a ver com as disposições sobre as multas pecuniárias do art. 1º §4º. Quanto às sanções, Martin Eifert explica que elas servem para o cumprimento do *compliance-system* requerido pela lei, e não pela punição por erros de bloqueio singulares. Nas palavras do catedrático:

Trata-se de um mal-entendido amplamente difundido no debate público que falhas ou erros no apagamento ou retirada de uma (única) postagem poderia resultar em uma multa pecuniária em desfavor do provedor. Esse não é o caso. Apenas uma falha sistemática é sancionada. (...) A tensão entre uma alta medida de liberdade organizacional para os afetados, a asseguaração de cumprimento de obrigações por meio de seu reforço por multa pecuniárias e, ao mesmo tempo, uma previsibilidade suficiente de sanções semelhantes de falhas organizacionais a casos evidentes.<sup>213</sup>

Portanto, a *NetzDG* estabelece infrações administrativas que podem variar de 50 milhões de euros para empresas, até 5 milhões de euros para pessoas que deixarem de cumprir sistematicamente as obrigações da lei, ou seja, em resumo, a norma pode vir a punir quem falha em produzir os relatórios ou o faz de maneira incorreta, além daquele que deixar de prover ou garantir o procedimento de reclamações, e, por fim, aquele que falha em supervisionar o tratamento de reclamações ou que deixa de oferecer orientação e treinamento conforme estabelece a lei.

---

<sup>213</sup> EIFERT, Martin. *A Lei Alemã para a melhoria da aplicação da lei nas redes sociais (NetzDG) e a regulação da plataforma*, op.cit. nota 104. p. 69.

## SÍNTESE OBJETIVA E CONCLUSÕES

Em relação à síntese objetiva, ou seja, recuperando os principais temas expostos no trabalho, vimos que o fenômeno global da desinformação nos obriga ao enfrentamento do problema sob a ótica da liberdade de expressão, mas também sob o ponto de vista democrático, pois o ecossistema informacional das Sociedades Algorítmicas vem operando de maneira disfuncional para a esfera pública. Como explicado, a personalização das redes, o micro-targeting e os modelos de negócios baseados em cliques acabaram alterando a lógica de consumo da informação, de maneira que um número razoavelmente extenso de pesquisadores deposita nessas características sistêmicas da arquitetura de controle das redes sociais a responsabilidade pela incapacidade dos indivíduos de se envolverem em discursos racionais e bem informados sobre a sociedade.

No entanto, se por um lado percebe-se que há diversos atores manipulando os debates públicos por meio do uso de contas automatizadas, havendo, inclusive, a ascensão de um modelo de guerra cibernética pela desinformação, por outro lado, trabalhos científicos mostram que o que de fato vem causando a disseminação de notícias falsas é o comportamento humano. Nesse sentido, ressalva-se que a compreensão científica de como e por que as *fake news* se espalham, baseia-se, atualmente, em análises pontuais, e não em larga escala. Assim, os primeiros desafios estão ligados a uma compreensão da desinformação como um *harm* sistêmico, que depende de uma maior colaboração científica a partir das perspectivas das empresas de tecnologia e da academia.

Em seguida, debatemos sobre as semelhanças e diferenças entre a noção clássica de esfera pública, qual seja, a concepção habermasiana de esfera pública burguesa, e a esfera pública digital das Sociedades Algorítmicas, com o intuito de mapear os impactos causados pelas redes sociais. Nessa toada, percebemos que a principal diferença entre os dois contextos históricos reside no fato de que, enquanto na esfera pública clássica tivemos o protagonismo da elite intelectual da época, hoje em dia qualquer

pessoa com um celular pode se tornar um produtor de conteúdo ou um influenciador político. Esse fenômeno pode ser interpretado pela ascensão do amadorismo nos meios de comunicação e pela simplificação dos debates políticos, não mais travados em jornais e revistas, mas em 140 caracteres.

Portanto, a premissa político-jurídica exposta até aqui consistiu no fato de que um ecossistema inundado de atores políticos que se utilizam da arquitetura das redes sociais para manipular os debates na esfera pública causa desinformação generalizada e isso é extremamente deletério para a democracia. Nessa toada, a conjuntura se mostra quase que um paradoxo em si mesma, pois a Internet, classicamente enxergada ao mesmo tempo como arena e ferramenta da liberdade de expressão, vem sufocando a cultura democrática que ela mesma visa fomentar, justamente pelo desdobramento da própria liberdade de expressão, sem freios e restrições, dentro do ecossistema informacional da atual Sociedade Algorítmica. A conclusão desta parte, portanto, é a de que a sociedade precisa fornecer algum tipo de resposta.

A partir desse momento, há um corte temático no trabalho, à medida que inserimos a desinformação e as *fake news* dentro da problemática mais abrangente dos principais *harms* praticados nas redes sociais. A razão para termos feito isso é que, para debatermos a regulação das redes sociais como um todo, devemos considerar que tanto as regulações para *fake news* quanto as regulações para outros *harms* - como o discurso de ódio, por exemplo - esbarrarão, igualmente, na liberdade de expressão. Sendo assim, antes de discutirmos salvaguardas específicas para cada *harm*, mostra-se necessário lidar com as dificuldades de regular setores disruptivos (com incertezas científicas e urgências sócio-políticas) e com a própria liberdade de expressão, a qual, como se sabe, desfruta de uma posição privilegiada nos ordenamentos constitucionais.

Passando para uma análise do panorama regulatório das novas intermediárias da Internet, vimos que o século XX apresentava um modelo dualista de regulação das intermediárias, com dois tipos básicos de atores: os Estados territoriais de um lado, e os indivíduos falantes do outro. Já o modelo regulatório do século XXI é pluralista, com múltiplos *players* e *stakeholders*, de maneira que seria mais fácil enxergar esse modelo regulatório como uma

relação triangular, havendo, atualmente, uma novidade na relação regulatória: a censura colateral feita pelas intermediárias nos discursos de seus usuários.

Resumidamente, essas plataformas interativas geralmente são baseadas em modelos de comunidade regulamentados por “termos de uso”, os quais cada usuário precisa aceitar e respeitar para ter acesso ao serviço. Assim, as plataformas ficam autorizadas a bloquear, censurar ou de outra forma controlar os discursos dos usuários, com base nos seus termos de uso ou políticas de serviço, os quais se inserem dentro de uma lógica de relação privada. No entanto, como elas baseiam sua receita nas mais variadas formas de captação de recursos pela exploração da sua principal *commodity*, a atenção dos usuários, as plataformas convivem com a constante necessidade de manter a boa governança de seus reinos digitais, sopesando técnicas que combinam censura de discursos feitos por humanos e máquinas, e as pressões mercadológicas pela expansão da rede social.

Neste ponto, explicamos que o desenvolvimento do mercado das novas tecnologias de comunicação aconteceu amparado em proteções legais e jurisprudências que permitiram a isenção de responsabilidade por divulgação de conteúdo ilícito (ou protegido) por terceiros. Sem essas exceções, seria economicamente impossível que essas empresas se desenvolvessem.

Em seguida, para adentrarmos ao debate regulatório, destacamos alguns comportamentos prejudiciais (*online harms*) praticados nas plataformas, frisando que a veiculação de *fake news* é apenas um dentre os vários comportamentos com lacunas legais. Aqui, abrimos tópico específico para comentar sobre a tendência internacional em responder a problemas digitais emergentes com a legislação já existente, minimizando-se, assim, a criação de novos tipos legais, os quais encontram dificuldades impostas pela arquitetura do código.

Em suma, as premissas que tentamos traçar nesta parte tem a ver com o fato de que, de um lado temos a dificuldade em regular tecnologias emergentes, que exigem ações multidisciplinares, e por isso devem estar embasadas em constatações científicas e técnicas de inúmeras áreas do conhecimento, o que demanda tempo. Por outro lado, temos urgência, pois os danos causados pelas e nas redes sociais são cada vez mais aparentes, já sendo certo que o comportamento humano está sendo modificado como

resultado das mudanças introduzidas pelas vidas online e pelas Sociedades Algorítmicas.

É nesse contexto de urgência social e de incertezas científicas que entram em cena mecanismos regulatórios que se propõem a antecipar e limitar as consequências prejudiciais à sociedade antes que elas ocorram - são mecanismos baseados no risco da atividade.

Por essa razão, pesquisadores vêm indicando que a melhor forma de lidar com essas dificuldades trazidas pelas tecnologias disruptivas é procurar respostas dentro das legislações já existentes e das possibilidades arquitetônicas da Internet. Em suma, qualquer resposta regulatória, precisa, em primeiro lugar, respeitar as possibilidades técnicas do setor; em segundo lugar, deve buscar estar embasada o máximo possível em leis já existentes, para ter que lidar o mínimo possível com a criação de novas espécies legais, que talvez não se adequem ao modelo arquitetônico (código). No entanto, para lidar com essas e outras dificuldades, é possível extrair alguns princípios importantes de outros regimes regulatórios, que sintetizam o norte das propostas regulatórias, como o princípio da precaução (*precautionary principle*) e o dever de cuidar (*duty to care*).

Para resumir, vislumbra-se um modelo regulatório em que a violação do dever legal pode levar à responsabilização das plataformas, mas no qual as questões de padrões a serem seguidos não são deixadas apenas nas mãos da indústria ou dos governos, de modo que haja transparência e que os procedimentos internos sigam regras de ordem pública, respeitando o contraditório das partes, diminuindo as chances de *overblocking*, e não penalizando as plataformas por questões impossíveis de serem resolvidas devido à arquitetura do código.

Por fim, partimos para análise da NetzDG. A lei alemã pode ser caracterizada com quatro objetivos principais. O primeiro objetivo da lei é melhorar o *enforcement* de legislações do mundo físico no mundo digital, ou seja, a lei está preocupada em melhorar a aplicação de legislação já existente no mundo digital. O segundo objetivo limita-se a obrigar as maiores plataformas a prestarem informações relativas aos ilícitos praticados na rede, as medidas tomadas para impedi-los, e a quantidade e o escopo das reclamações enviadas às empresas pelos usuários. Com isso, a lei busca aumentar a transparência

dessas empresas. O terceiro objetivo consiste em regulamentar os processos internos das empresas no que tange ao tratamento de reclamações sobre a prática desses ilícitos nas redes sociais, sendo este o objetivo mais complexo e criticado, porque contém a obrigação de retirar ou bloquear conteúdo evidente e manifestamente ilícito do ar em 24 horas ou em 7 dias para casos menos óbvios. A crítica veiculada aqui tem a ver com a aplicação do Direito pelas empresas de tecnologia. O quarto objetivo principal da lei foi estabelecer sanções administrativas e pecuniárias às empresas que descumprem as obrigações legais, podendo as sanções variarem de 5 milhões para pessoas, a 50 milhões de euros para empresas.



## COMENTÁRIOS METODOLÓGICOS E PROCEDIMENTAIS

Do ponto de vista metodológico, é possível sintetizar os temas do trabalho em três tópicos: (i) a produção da desinformação no ecossistema informacional nas esferas públicas das democracias conectadas; (ii) as consequências político-democráticas e sociais da desinformação; e (iii) as possíveis soluções estatais e não estatais de combate à desinformação.

Embora a centralidade e atualidade do tema dificultem a delimitação exata da questão formulada, é admissível considerarmos como problema colocado o fato de as democracias modernas estarem vivenciando distorções informacionais, bem como a polarização e o extremismo político, ocasionados pela manipulação de debates - e, conseqüentemente, de eleições - por meio de uma arquitetura de controle natural às novas plataformas de comunicação, mas que precisam sofrer alterações em nome da maximização de princípios político-democráticos.

Quanto ao método, embora não seja possível dizer que este trabalho esteja embasado no método hipotético-dedutivo, de modo que não será produzida uma hipótese e tentativa de falseamento, para fins acadêmicos a resposta (provável e provisória) que o trabalho busca confirmar (ou negar) é a necessidade (ou desnecessidade) de regulação da liberdade de expressão nas plataformas interativas, o que, de certa forma, se apresenta como uma hipótese. Portanto, para fins de averiguação, a pergunta de partida seria: *“diante do cenário de desinformação, é necessária a regulação da liberdade de expressão nas plataformas interativas no Brasil?”*. E a resposta, mesmo sendo notadamente inconclusiva, seria: *“Sim, o cenário brasileiro demanda regulações brandas no setor, porque o desenvolvimento da liberdade de expressão, sem freios e restrições, dentro das plataformas interativas vêm causando, além da desinformação, uma série de outros danos a usuários, esvaziando a esfera pública e colocando em risco a própria liberdade de expressão e o ideal democrático que ela visa proteger.”*

Acerca da pesquisa, considerando que o trabalho conta com a análise das posições norte-americana e alemã, com a finalidade de situar o Brasil no debate, a dissertação se pauta também no método do estudo de ordenamentos internacionais - os quais podem ser considerados como casos estudados -, mas também sob uma perspectiva comparativa da situação nestes países e os respectivos ordenamentos com o brasileiro. Neste ponto, acredita-se que foi cumprida a função do uso do método do Direito Comparado, pois, em relação ao tema estudado, ele serviu para tornar evidentes as diferenças e semelhanças existentes entre regulações jurídicas de sistemas jurídicos distintos, além de ter contribuído para a propositura de eventuais reformas legislativas e judiciais - desde que adequadas à realidade brasileira. É interessante notar que o Brasil ainda busca se adequar à nova realidade comunicacional apresentada, ao passo que outros países, como a Alemanha, recentemente desenvolveram regras específicas visando solucionar problemas concretos.

Do ponto de vista da coleta de material, uma substancial parte do trabalho foi desenvolvida com pesquisa bibliográfica na rede mundial de computadores, a partir do acesso aos sites dos principais centros de estudo de Internet e Sociedade ao redor do mundo, havendo uma concentração nas Universidades de Harvard, Yale, Columbia, Humboldt e Oxford. Não obstante a utilização de literatura internacional, também foram analisados escritos de diversos pesquisadores brasileiros, bem como livros e artigos nacionais, devidamente contidos nas referências. Além disso, orientações, indicações e *brainstormings* fundamentais para o desenvolvimento do trabalho partiram de aulas, palestras e encontros ocorridos tanto na Universidade do Estado do Rio de Janeiro (UERJ), quanto na Procuradoria do Estado do Rio de Janeiro (PGE-RJ).

Nesse contexto, partimos das considerações desenvolvidas por pensadores-chaves, como Cass Sunstein, Jack Balkin, Tim Wu, Karl-Heinz Lauer, Martin Eifert, Wolfgang Schulz, e Thomas Vestig, bem como diversos pesquisadores nacionais que vêm desenvolvendo substanciais estudos sobre o tema, como Carlos Affonso Souza, Nelson Nery Junior, Ricardo Campos, Ronaldo Porto Macedo Junior, Ronaldo Lemos, Sérgio Branco, Eduardo Magrani, Diogo Rais, entre outros.

Identificou-se que, apesar de certa similitude entre os casos estudados (e outros a serem estudados futuramente, mas que por questões temporais se mostraram inviáveis, como o caso da recente lei francesa regulando o uso das redes sociais em momentos eleitorais), há elementos e argumentos centrais, problemas levantados, e hipóteses formuladas, comuns e identificáveis entre os trabalhos pesquisados, embora haja singularidades nos posicionamentos de cada pesquisador.

Tendo em vista tratar-se de estudos disciplinarmente localizados no âmbito do Direito Constitucional, Administrativo-Regulatório, Eleitoral e Digital, ou melhor, nas interseções dos estudos de Direito e Democracia, as observações aqui feitas podem auxiliar no desenvolvimento de eventuais normas que regularão o uso das principais redes sociais utilizadas no Brasil, contribuindo possivelmente para um ecossistema informacional mais saudável.

Neste ponto, destaca-se que, se de um lado o trabalho está amparado em posicionamentos e constatações comuns notados e documentados ao redor do mundo, de outro lado também está amparado em algumas incertezas teóricas e científicas, naturais aos temas tecnológicos e aos debates sobre comportamento humano, especialmente se considerarmos a novidade do tema e a dificuldade de experimentação empírica. Além disso, debater temas tangentes às redes sociais esbarra na problemática do acesso a informações empresariais sigilosas, como os dados das empresas que gerem redes sociais. Tal fato é relatado em quase todas as pesquisas sobre redes sociais, que chegam a chamar os algoritmos de *black boxes*.

Não obstante existirem incertezas, a alta incidência de argumentos e posicionamentos comuns nas opiniões dos principais pesquisadores do tema permite organizar um arcabouço teórico suficientemente capaz de fundamentar respostas às questões propostas e de nos auxiliar a compreender como os elementos teóricos se colocam na prática, bem como em que medida ocorre uma resignificação desses elementos a depender da sociedade analisada. Os conceitos-chaves que serão analisados neste trabalho podem ser divididos em: (a) a democracia, havendo um enfoque para o modelo deliberativo; (b) a esfera pública clássica e a esfera pública digital, notando-se a construção contestatória da conceptualização habermasiana de esfera pública burguesa; (c) a liberdade de expressão no contexto da Internet, recortada a

partir da *public forum doctrine* e da evolução jurisprudencial da Primeira Emenda norte-americana; (d) o panorama regulatório das intermediárias da comunicação, destacando-se a hibridez dos novos intermediários; e (e) a autorregulação regulada das plataformas, observada a partir da nova lei alemã para melhoria de aplicação das leis nas redes sociais (*NetzDG*).

Justamente a partir da análise dos conceitos supracitados, também sob a perspectiva do ordenamento brasileiro, coube situar o Brasil na discussão da regulação dos discursos nas redes sociais. Para tal, há um olhar para as decisões jurisprudências e para a legislação temática, como o Marco Civil da Internet e a Lei Geral de Proteção de Dados, entre outras.

Finalmente, no tocante aos objetivos, estabeleceu-se o objetivo geral de identificar como é produzida a desinformação no contexto do consumo de informação via redes sociais, apontando as principais mudanças e consequências desse novo ecossistema informacional, bem como indicar possíveis soluções.

Quanto aos objetivos específicos, pode-se dizer que se aproximam da tentativa de, em primeiro lugar, destrinchar os diferentes tipos de desinformação existentes, além dos modelos de propaganda e arquitetura das intermediárias da comunicação, destacando-se algumas mazelas para as democracias daí advindas. Em segundo lugar, objetivou-se diferenciar a esfera pública clássica e as restrições impostas ao direito de liberdade de expressão de períodos pré-redes sociais, por um lado, da esfera pública digital, por outro, debatendo as possíveis restrições à fala online, ou seja, eventuais regulações constitucionalmente viáveis. Como terceiro objetivo específico, buscou-se dividir as possíveis soluções em dois pilares: as estatais e as não estatais.

## REFERÊNCIAS

ABBOUD, Georges; Campos, Ricardo. *A autorregulação regulada como modelo do Direito Proceduralizado: regulação de redes sociais e proceduralização*. in in Abboud, Georges. Nery Jr., Nelson. e Campos, Ricardo (Coordenadores) . São Paulo: Thomson Reuters Brasil, 2018.

BALKIN, Jack M. *Cultural Democracy and the First Amendment*. 110 Nw. U. L. Rev. 1053 (2016). Disponível em <<https://scholarlycommons.law.northwestern.edu/nulr/vol110/iss5/3>>

BALKIN, Jack M. *Free Speech in the Algorithmic Society: Bid Data, Private Governance, and New School Speech Regulation*. Disponível em: <[https://lawreview.law.ucdavis.edu/issues/51/3/Essays/51-3\\_Balkin.pdf](https://lawreview.law.ucdavis.edu/issues/51/3/Essays/51-3_Balkin.pdf)>

BALKIN, Jack M., "Old-School/New-School Speech Regulation" (2014). Faculty Scholarship Series. 4877. Disponível em <[https://digitalcommons.law.yale.edu/fss\\_papers/4877](https://digitalcommons.law.yale.edu/fss_papers/4877)>

BALKIN, Jack M., *Digital Speech and Democratic Culture: A Theory of Freedom of Expression for the Information Society* (2004). Faculty Scholarship Series. 240. p. 31-32. Disponível em <[https://digitalcommons.law.yale.edu/fss\\_papers/240](https://digitalcommons.law.yale.edu/fss_papers/240)>

BALKIN, Jack M., *Free Speech is a Triangle*. Columbia Law Review, 2018, Yale Law School, Public Law Research Paper N. 640. Disponível em: <<https://ssrn.com/abstract=3186205>>

BALKIN, Jack M., *The Three Laws of Robotics in the Age of Big Data* (August 27, 2017). Ohio State Law Journal, Vol. 78, (2017), Yale Law School, Public Law Research Paper No. 592. Disponível em: <<https://ssrn.com/abstract=2890965>>

BAPTISTA, Patrícia; Keller, Clara Iglesias. *Por que, quando e como regular as novas tecnologias? Os desafios trazidos pelas inovações disruptivas*. Revista de Direito Administrativo, Rio de Janeiro, v. 273, p. 123-163, set. 2016. ISSN 2238-5177. Disponível em: <<http://bibliotecadigital.fgv.br/ojs/index.php/rda/article/view/66659>>

BENHABIB, Seyla. "Models of Public Space: Hannah Arendt, the Liberal Tradition, and Jürgen Habermas" (p. 73-98)

BENHABIB, Seyla. *Sobre um modelo deliberativo de legitimidade democrática*. In: Melo, Rúrion Soares; Werle, Denílson Luiz (Eds.). Democracia Deliberativa. São Paulo: Editora Singular, 2007, p. 47-79.

BINENBOJM, Gustavo. *Poder de polícia, ordenação, regulação: transformações político-jurídicas, econômicas e institucionais do direito administrativo ordenador*. 2. ed. Belo Horizonte: Fórum, 2017, p.143-153.

CALHOUN, Craig (ed.). *Habermas and the Public Sphere*. Cambridge, MA: MIT Press, (1992).

CAMPOS, Ricardo. Maranhão, Juliano. *Fake news e autorregulação regulada das redes sociais no Brasil: fundamentos constitucionais*. in Abboud, Georges. Nery Jr., Nelson. e Campos, Ricardo (Coordenadores). *Fake News e Regulação*. Coleção Direito do Estado e Transformação. São Paulo: Thomson Reuters Brasil, 2018

CHRISTIANO, Thomas. *A importância da deliberação pública*. In: Melo, Rúrion Soares; Werle, Denílson Luiz (Eds.). *Democracia Deliberativa*. São Paulo: Editora Singular, 2007, p. 81-113.

CHRISTIANO, Thomas. *Procedimento e substância na democracia deliberativa*. In: MELO, Rúrion Soares; WERLE, Denílson Luiz (Eds.). *Democracia Deliberativa*. São Paulo: Editora Singular, 2007, p. 115-144.

COHEN, Joshua. *Deliberation and democratic legitimacy*. In: Hamlin, Alan; Pettit, Philip (Eds.). *The Good Polity*. Oxford: Blackwell, 1989, p. 17-34

CONRAD, Albrecht. *Exceptions to copyright protection and the permitted uses of copyright works in the hi-tech and digital sectors*. Disponível em: <[https://www.aippi.org/download/comitees/216/GR216germany\\_en.pdf](https://www.aippi.org/download/comitees/216/GR216germany_en.pdf)>(acesado em 04.02.19).

D I KRAMER, Adam & Guillory, Jamie & Hancock, Jeffrey. *Experimental Evidence of Massive-Scale Emotional Contagion Through Social Networks*. Proceedings of the National Academy of Sciences of the United States of America. 111. 10.1073/pnas.1320040111. Disponível em <[https://www.researchgate.net/publication/262813340\\_Experimental\\_Evidence\\_of\\_Massive-Scale\\_Emotional\\_Contagion\\_Through\\_Social\\_Networks](https://www.researchgate.net/publication/262813340_Experimental_Evidence_of_Massive-Scale_Emotional_Contagion_Through_Social_Networks)>

DAVID M. J. Lazer, Matthew A. Baum, Yochai Benkler, Adam J. Berinsky, Kelly M. Greenhill, Filippo Menczer, Miriam J. Metzger, Brendan Nyhan, Gordon Pennycook, David Rothschild, Michael Schudson, Steven A. Sloman, Cass R. Sunstein, Emily A. Thorson, Duncan J. Watts & Jonathan L. Zittrain, *The Science of Fake News*, *Science*, 9, Mar., 2018.

DENHAM, Elizabeth. *The future of Internet Harms Regulation*. Paper for the Digital, Culture, Media and Sport Committee. Supplementary written evidence submitted by Elizabeth Denham, Information Commissioner. Disponível em <<http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/digital-culture-media-and-sport-committee/disinformation-and-fake-news/written/92853.html>>

EIFERT, Martin. *A Lei Alemã para a melhoria da aplicação da lei nas redes sociais (NetzDG) e a regulação da plataforma*. in Abboud, Georges. Nery Jr., Nelson. e Campos, Ricardo (Coordenadores). *Fake News e Regulação*. Coleção Direito do Estado e Transformação. São Paulo: Thomson Reuters Brasil, 2018

FARIS, Robert and Ashar, Amar and Gasser, Urs and Joo, Daisy, *Understanding Harmful Speech Online*. Berkman Klein Center Research Publication N°. 2016-2. Disponível em: <<https://ssrn.com/abstract=2882824>>

FRASER, Nancy. "Rethinking the Public Sphere: A Contribution to the Critique of Actually Existing Democracy" (p. 109-142).

GELFERT, Axel. *Fake News: A Definition*. *Informal Logic* 38 (1) (2018): 84-117.

GIANCCHETTA, André Zonaro. *Atuação e responsabilidade dos provedores diante das fake news e da desinformação*. p.27. in *Fake News: a conexão entre a desinformação e o direito*. Coord. Rais, Diogo. São Paulo: Thomson Reuters Brasil, 2018.

GOLLATZ, Kirsten. Riedl, Martin J. Pohlmann, Jens. *Removals of online hate speech in numbers*. Disponível em: <<https://www.hiig.de/en/removals-of-online-hate-speech-numbers/>>

GUTWIRTH, Serge. De Hert, Paul. De Sutter, Laurent. *The trouble with technology regulation from a legal perspective. Why Lessig's 'optimal mix' will not work*. Oxford Regulating Technologies. Legal Futures, Regulatory Frames and Technological Fixes (2008). Disponível em [http://works.bepress.com/serge\\_gutwirth/1/](http://works.bepress.com/serge_gutwirth/1/)

HABERMAS, Jürgen. "Espaço público e esfera política pública. Raízes biográficas de dois motivos de pensamento", in: *Entre Naturalismo e Religião*. Rio de Janeiro: Tempo Brasileiro, 2007, p. 17-30

HABERMAS, Jürgen. "Further Reflections on the Public Sphere", in: Calhoun, Craig (ed.). *Habermas and the Public Sphere*, op. cit., p. 421-461)

HABERMAS, Jürgen. "Political communication in media society: Does democracy still enjoy an epistemic dimension? The impact of normative theory on empirical research", *Communication Theory*, 16 (2006): 411-426.

HABERMAS, Jürgen. "Religion in the public sphere", *European Journal of Philosophy*, 14 (2006)

HABERMAS, Jürgen. *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy*. Cambridge, MA: MIT Press, 1996.

HABERMAS, Jürgen. *Mudança Estrutural da Esfera Pública*. São Paulo: Editora UNESP, 2014,

HABERMAS, Jürgen. *Strukturwandel der Öffentlichkeit*. Neuwied/Berlim: Luchterhand, 1962 (tradução inglesa: *The Structural Transformation of the Public Sphere*. Cambridge, MA: MIT Press, 1989

KELLER, Daphne. *Empirical Evidence of "Over-Removal" By Internet Companies Under Intermediary Liability Laws*, The Center for Internet and Society at Stanford Law School, (2017). Disponível em <<http://cyberlaw.stanford.edu/blog/2015/10/empirical-evidence-over-removal-internet-companies-under-intermediary-liability-laws>>

KOSINSKI, Michal. Stillwell, David. Graepel, Thore. *Digital records of behavior expose personal traits*. Proceedings of the National Academy of Sciences Apr

2013, 110 (15) 5802-5805; DOI: 10.1073/pnas.1218772110. Disponível em: <<http://www.pnas.org/content/110/15/5802#ref-15>>

LADEUR, Karl-Heinz. *Por um novo direito das redes digitais: digitalização como objeto contratual, uso contratual de 'meios sociais', proteção de terceiros contra violações a direitos da personalidade por meio de Cyber Courts*. in Abboud, Georges. Nery Jr., Nelson. e Campos, Ricardo (Coordenadores). *Fake News e Regulação*. Coleção Direito do Estado e Transformação. São Paulo: Thomson Reuters Brasil, 2018

LESSIG, Lawrence. *Code. Version 2.0*. New York: Basic Books. 2006.

LIDSKY, Lyrisa B. *Government Sponsored Social Media and Public Forum Doctrine under the First Amendment: Perils and Pitfalls*, 19 Pub.Law. 2 (2011). Disponível em: <<https://scholarship.law.ufl.edu/cgi/viewcontent.cgi?article=1641&context=facultypub>>

MACKINNON, Rebecca. Hickok, Elonnai. Bar, Allon. Lim, Hae-in. *Fostering Freedom Online: The role of Internet Intermediaries*. Series on Internet Freedom. Paris: United Nations Educational, Scientific and Cultural Organization (UNESCO), 2014

MAGRANI, Eduardo. *Democracia conectada: a internet como ferramenta de engajamento político-democrático*. Curitiba: Juruá, 2014

MAHLOULY, Dounia. "Rethinking the Public Sphere in a Digital Environment: Similarities between the Eighteenth and the Twenty-First Centuries, op.cit., p. 18.

MANIN, Bernard. *Legitimidade e deliberação política*. In: Melo, Rúrion Soares; Werle, Denílson Luiz (Eds.). *Democracia Deliberativa*. São Paulo: Editora Singular, 2007, p.15-45.

MENDONÇA, José Vicente Santos de. *Direito Constitucional Econômico: a intervenção do Estado na economia à luz da razão pública e do pragmatismo*. 2. ed. Belo Horizonte: Fórum, 2017, p. 359.

NABIHA Syed, *Real Talk About Fake News: Towards a Better Theory for Platform Governance*, 127 Yale L.J. F. 337 (2017). Disponível em <<http://www.yalelawjournal.org/forum/real-talk-about-fake-news>>

O'NEILL, Kevin F. *Disentangling the Law of Public Protest*, 45 Loyola Law Review 411 (1999).

PERSET, Karine. *The Economic and Social Role of Internet Intermediaries*. Paris, Organisation for Economic and Co-operation and Development/OECD. March 2010, Disponível em <[www.oecd.org/internet/ieconomy/44949023.pdf](http://www.oecd.org/internet/ieconomy/44949023.pdf)>

PERSILY, Nathaniel. *Can Democracy Survive the Internet?*, 28 Journal of Democracy 63 (2017).

RAUCHFLEISCH, Adrian. "The public sphere as an essentially contested concept: A co-citation analysis of the last 20 years of public sphere research". *Communication and the Public*, 2/1 (2017): 3-18



RAWLS, John. *O Direito dos Povos*. São Paulo: Martins Fontes, 2001, p. 174-175

SCHULZ, Wolfgang, Regulating Intermediaries to Protect Privacy Online – The Case of the German NetzDG (July 19, 2018). Marion Albers and Ingo Sarlet (editors), *Personality and Data Protection Rights on the Internet*. Disponível em <<https://www.hiig.de/wp-content/uploads/2018/07/SSRN-id3216572.pdf>>

SOROUGH, Vosoughi. Deb, Roy. Sinan, Aral. The spread of true and false news online. *Science*, v. 359, 9, Mar. 2018, pp.1146-1151. Disponível em: <<http://science.sciencemag.org/content/359/6380/1146.full>>

SOUZA, Carlos Affonso. Lemos, Ronaldo. *Marco civil da internet: construção e aplicação*. Juiz de Fora: Editar Editora Associada Ltda, 2016. Disponível em <[https://itsrio.org/wp-content/uploads/2017/02/marco\\_civil\\_construcao\\_aplicacao.pdf](https://itsrio.org/wp-content/uploads/2017/02/marco_civil_construcao_aplicacao.pdf)>

SOUZA, Carlos Affonso. Padrão, Vinicius. Quem lê tanta notícia (falsa)? Entendendo o combate contra as “fake news”. Disponível em: <<https://feed.itsrio.org/quem-l%C3%AA-tanta-not%C3%ADcia-falsa-entendendo-o-combate-contra-as-fake-news-70fa0db05aa5>>

SUNSTEIN, Cass. R. *Human Agency and Behavioral Economics*. Cham, Switzerland: Palgrave Macmillan, 2017.

SUNSTEIN, Cass. R. *#Republic: Divided Democracy in the Age of Social Media*. Princeton Univ. Press. Mar. 2017

THALER, R. H., & Sunstein, C. R. (2009). *Nudge: improving decisions about health, wealth, and happiness*. Rev. and expanded ed. New York: Penguin Books

WAACK, William. Fake News: uma visão político-jornalística. in Abboud, Georges. Nery Jr., Nelson. e Campos, Ricardo (Coordenadores). *Fake News e Regulação*. Coleção Direito do Estado e Transformação. São Paulo: Thomson Reuters Brasil, 2018.

WU, Tim. *Is the First Amendment Obsolete?* Columbia Public Law Research Paper No. 14-573. (2017). p. 6. Disponível em: <<https://ssrn.com/abstract=3096337>>