

## A esfera pública (forjada) na era das fake news e dos filtros-bolha

EDUARDO MAGRANI  
RENAN MEDEIROS DE OLIVEIRA

### RESUMO

■ No presente artigo pretendemos explorar, através da revisão bibliográfica e do estudo de pesquisas de intenção de votos, um pouco dos novos fenômenos tecnológicos que, juntos, impactam a forma como o cidadão forma sua convicção acerca dos fatos cotidianos importantes para a vida pública, de modo geral, e do processo eleitoral e dos candidatos, de modo particular. Em primeiro lugar, fazemos uma breve abordagem do arcabouço teórico em que nos baseamos para pensar numa esfera pública comunicativa, racional e na qual se busca a situação ideal de fala. Em seguida, tratamos das fake news – que se tratam de notícias falsas que buscam influenciar a forma como a população olha para determinado candidato – e das *deep fakes* – as quais têm objetivo similar, mas agem por meio da alteração da realidade de forma mais profunda. Por fim, abordamos como os algoritmos, sobretudo o uso de *bots*, estão agindo de modo a criar uma esfera pública forjada, que não condiz com o real desejo e com a real necessidade dos indivíduos. Além disso, tratamos de como o pensamento dos indivíduos está sendo distorcido no cenário das *filter bubble*, que potencializam os efeitos dos fenômenos estudados nos itens precedentes. Buscaremos demonstrar, ao longo do desenvolvimento deste estudo e através do método hipotético-dedutivo, que as novas tecnologias possuem um grande potencial de impacto na vontade eleitoral, por mais que esse potencial ainda não tenha sido explorado em toda sua extensão. Caminha-se para um cenário em que o processo eleitoral é hackeável.

## ABSTRACT

■ In the present article, we intend to explore, through the bibliographical review and the study of vow intention research, some of the new technological phenomena that, together, impact the way the citizen forms his conviction about the daily facts important for public life, in general, and of the electoral process and the candidates, in particular. In the first place, we make a brief approach to the theoretical framework on which we base ourselves to think of a communicative, rational public sphere and in which the ideal speech situation is sought. Next, we deal with the fake news – false news that seeks to influence the way the population looks at a particular candidate – and the deep fakes – which have a similar purpose but act by changing reality more deeply. Finally, we discuss how the algorithms, especially the use of bots, are acting to create a forged public sphere that does not match the real desire and the real need of the individuals. In addition, we deal with how the thinking of individuals is being distorted in the filter bubble scenario, which potentiates the effects of the phenomena studied in the previous items. Throughout the development of this study and through the hypothetico-deductive method, we will try to demonstrate that the new technologies have a great potential of impact in the electoral will, although this potential has not yet been explored in all its extension. He walks to a scenario where the electoral process is hackable.

## INTRODUÇÃO

■ As fake news já se mostraram potentes influenciadoras do processo eleitoral. No momento de formar sua opinião, o eleitor sofre o impacto de notícias cuja veracidade não é averiguada, criando um juízo em relação aos candidatos e ao processo democrático baseado em notícias falsas. Ainda não é possível afirmar a exata dimensão exercida pelas fake news no processo eleitoral, mas é fato que alguma influência elas exercem.

A potencial prejudicialidade das fake news é exponenciada quando consideramos a forma como as novas tecnologias estão sendo utilizadas em conjunto. As *deep fakes*, os algoritmos, o *filter bubble*, dentre outros, colocam em xeque a forma como se olha para a realidade, afetando aspectos da vida que vão além das eleições. Questionar o *status quo* e a veracidade de acontecimentos é algo positivo e necessário em uma democracia. Contudo, é preciso haver consensos mínimos sobre os fatos, sobretudo os de interesse público. O grande volume de notícias

que colocam em dúvida a maneira como as coisas se deram na realidade diminui a capacidade das pessoas de diferenciar o real do inventado (Viana, Zanatta, 2018). É indispensável que padrões éticos básicos sejam respeitados com o intuito de se assegurar um ambiente democrático minimamente saudável.

O cenário se agrava quando se leva em consideração que a mídia tradicional, sobretudo a televisão, está perdendo espaço e confiança. O cidadão não acredita mais em tudo o que é dito na TV, taxando os conteúdos de tendenciosos e fora de contexto. A televisão ainda exerce, sim, papel fundamental, mas é preciso ter em conta que esse papel vem sendo reduzido e dando espaço para a internet, com foco nas redes sociais<sup>1</sup>. Contudo, por mais que a internet seja uma fonte de inesgotável conteúdo e possibilite a busca por informações por parte do usuário, o fenômeno que ficou conhecido como *filter bubble* cria empecilhos para que um ambiente dialógico *online* saudável e democraticamente desejável.

No presente artigo pretendemos explorar um pouco dos fenômenos que, juntos, impactam a forma como o cidadão forma sua convicção acerca dos fatos cotidianos importantes para a vida pública, de modo geral, e do processo eleitoral e dos candidatos, de modo particular. Em primeiro lugar, fazemos uma breve abordagem do arcabouço teórico em que nos baseamos para pensar numa esfera pública comunicativa, racional e na qual se busca a situação ideal de fala. Em seguida, tratamos das fake news e das *deep fakes*. Em poucas palavras, aquelas se tratam de notícias falsas que buscam influenciar a forma como a população olha para determinado candidato. As *deep fakes* têm objetivo similar, mas agem por meio da alteração da realidade de forma mais profunda. Por fim, abordamos como os algoritmos, sobretudo o uso de *bots*, estão agindo de modo a criar uma esfera pública forjada, que não condiz com o real desejo e com a real necessidade dos indivíduos. Além disso, tratamos de como o pensamento dos indivíduos está sendo distorcido no cenário das *filter bubble*, que potencializam os efeitos dos fenômenos estudados nos itens precedentes.

Para os fins aqui buscados, nos valeremos, sobretudo, da revisão bibliográfica sobre fake news, *deep fakes*, *bots* e *filter bubble* e sobre o impacto desses fenômenos nas eleições e na formação de convicção dos indivíduos. Também recorreremos ao levantamento de pesquisas de intenção de votos. Assim, buscaremos demons-

---

1 Os dados mostram uma clara distinção geracional em relação às fontes de obtenção de informação: quanto mais elevada a idade, maior o uso da televisão como principal veículo de comunicação. Até os 24 anos, por sua vez, mais da metade dos jovens utiliza a internet como principal meio. Confira-se os dados da Pesquisa Brasileira de Mídia 2016, disponíveis em: <<http://pesquisademidia.gov.br/#/Geral>>. Acesso em: 29 out. 2016.

trar, ao longo do desenvolvimento deste estudo e através do método hipotético-dedutivo, que as novas tecnologias possuem um grande potencial de impacto na vontade eleitoral, por mais que esse potencial ainda não tenha sido explorado em toda sua extensão. Caminha-se para um cenário em que o processo eleitoral é hackeável.

## I. BREVE NOTA TEÓRICA: A ESFERA PÚBLICA VIRTUAL

■ Um dos intuitos deste estudo é apontar a necessidade de existirem padrões éticos mínimos no uso das novas tecnologias e mecanismos para contornar os abusos advindos da perspectiva utilitarista, impedindo que se use uma pessoa como um meio e não como um fim em si mesma. Com isso, queremos impedir formas de manipulação de perfis reais ou a utilização de *bots* com o intuito de criar prioridades forjadas na agenda pública. Podemos pensar qual a perspectiva ética mais adequada para lidar com a tecnologia em um contexto em que os procedimentos e as ações democráticas estão relacionados ao complexo mundo de dados e de constante interação homem-máquina (coisa) em que vivemos. Impõe-se, portanto, um ‘dever ser’ ético e moral não somente atrelado às finalidades, mas também a todo o procedimento e gama de ações

Para isso, entendemos que é preciso levar em conta a perspectiva teórica completa e complexa de Jürgen Habermas, a qual nos permite pensar sobre o avanço deste novo mundo de dados de maneira dialógica e participativa para atingir proposições regulatórias mais legítimas e consensuais<sup>2</sup>.

O pensador alemão, nascido em 1929, vivenciou na Alemanha pós-guerra, com os julgamentos de Nuremberg, a profundidade do fracasso moral e político da Alemanha no âmbito do nacional-socialismo (Bohman, Rehg, 2014). Habermas destacou-se no mundo acadêmico ao analisar o desenvolvimento da esfera pública burguesa desde as origens, nos salões do século XVIII, até a sua transformação através da influência de meios de comunicação dirigidos pelo capital<sup>3</sup>.

Para Habermas, a legitimidade das normas e do sistema político em sociedades ocidentais capitalistas contemporâneas depende da aceitação das normas pelos cidadãos (Habermas, 2003, p. 16). Isso ocorre por meio de sucessivas tentati-

---

2 Os principais conceitos e formulações de Jürgen Habermas e sua relação com as plataformas de internet podem ser conferidos em estudo de Eduardo Magrani (2014).

3 Com a publicação em 1962 de sua habilitação, *Strukturwandel der Öffentlichkeit* (Transformação Estrutural da Esfera Pública, ed. Inglesa, 1989).

vas de justificação nas quais cada cidadão deve vincular livremente sua vontade ao conteúdo da norma através de um processo racional e dialógico de argumentação, isto é, de reflexão e convencimento (cf. Cohen, 1997, p. 29).

Neste tipo de sociedade, a esfera pública é entendida justamente como o conjunto de espaços que permitem a ocorrência dos processos dialógicos comunicacionais de articulação de opiniões e de reconstruções reflexivas dos valores, disposições morais e normativas que orientam a convivência social. É na esfera pública que os diferentes grupos constitutivos de uma sociedade múltipla e diversa partilham argumentos, formulam consensos e constroem problemas e soluções comuns (Magrani, 2014).

A esfera pública de Habermas constitui uma zona de intercâmbio entre, de um lado, o sistema – caracterizado como o mundo do trabalho, pautado pela lógica do dinheiro e do poder, como um mundo instrumental de ação estratégica, e não comunicativa, orientado pelo mercado e pela burocracia (Habermas, 1987, p. 113-197; Calhoun, 1992, p. 1-51) –, e, de outro lado, os espaços públicos e privados do mundo da vida – caracterizado como o mundo da interação entre pessoas, que se organizam comunicativamente através da língua ordinária viabilizando a ação comunicativa sem um agir estratégico, orientada somente para o entendimento intersubjetivo que conduz idealmente ao acordo ou leva ao consenso (Habermas, 2003, p. 107).

Ao analisar a transformação da esfera pública burguesa dirigida pelo poder do capital por meio da influência de meios de comunicação, Habermas alerta para a forte tendência à “colonização do mundo da vida” pelo sistema e seus valores. A colonização decorreria da intromissão da política e economia no mundo da vida, responsável pela redução da cidadania e transformação dos cidadãos em clientes dos serviços de bem-estar social, sendo esta a marca da modernidade. Neste cenário, o poder do capital econômico e da política invadem destrutivamente o mundo da vida. Segundo Habermas, a intervenção sistêmica impacta destrutivamente a reprodução cultural, na integração social e na socialização, como componentes do mundo da vida<sup>4</sup>.

Embora o autor não tenha se debruçado específica e deliberadamente sobre o tema da internet, defendemos a possibilidade de se compreender as plataformas digitais como esferas públicas abstratas, dotadas de grande potencial

---

4 Apesar de Habermas prever que não haja uma blindagem completa do mundo da vida da lógica sistêmica, acredita na capacidade dessa lógica ser anulada pela própria dinâmica do mundo da vida, pautado no agir comunicativo.

comunicativo e democrático<sup>5-6</sup>. Encontramos nos espaços digitais conectados uma esfera pública na qual indivíduos se comunicam regularmente, através de fóruns de discussão, redes sociais, ou plataformas de troca de mensagens que se aproximam muito da concepção de esfera pública desenhada por Habermas em menor escala.

No entanto, com o avanço das tecnologias digitais mais recentes, acompanhamos a transformação também desses espaços conectados, sendo possível vislumbrar uma possível redução no seu potencial comunicativo democrático.

Observamos hoje a predominância nas esferas conectadas dos lucrativos modelos de negócio baseados em filtragem algorítmica com a finalidade de realizar práticas de *microtargeting*, *profiling*, entre outras, direcionando a venda de produtos e serviços de forma otimizada a *e*-consumidores. Essas práticas correntes pautam-se pela utilização em grande medida dos dados pessoais dos usuários e geram o agravamento do efeito denominado “*filter bubble*”, possuindo efeitos nocivos sobre a democracia e freando o entusiasmo acerca do papel democrático da internet como esfera pública para as sociedades contemporâneas<sup>7</sup>. Nos itens a seguir, estudamos alguns desses mecanismos e suas implicações éticas para o contexto democrático como um todo e para as eleições, de forma específica.

---

5 Para um tratamento aprofundado dessa defesa, cf. Magrani, 2014, p. 25 e ss.).

6 A teoria habermasiana isoladamente não nos auxilia de forma suficiente no aprofundamento das possíveis soluções para esses problemas, pois foi pensada principalmente para medir e induzir o comportamento do agente humano racional e dialógico que interage na esfera pública. Contudo, ela serve como um excelente paradigma para analisar as reais possibilidades de construção de um cenário de diálogo e de fala no contexto *online*.

7 Não há dúvidas hoje sobre a existência do efeito *filter-bubble* na esfera pública conectada. Com relação, no entanto, há escala e impacto desse efeito, encontramos opiniões distintas e as pesquisas nesse tema ainda são embrionárias. Além disso, a esfera pública conectada é altamente dinâmica, com seus algoritmos mudando constantemente, alterando o modo de funcionamento dos espaços de diálogo digital. Para Pablo Ortellado (2018), baseando-se em estudo recente sobre o tema (“Avoiding the Echo Chamber about Echo Chambers, Knight Foundation”, 2018): “O perigo dos guetos informacionais nas mídias sociais tem sido bastante superestimado. a polarização é um fenômeno circunscrito aos mais engajados, que são também os mais visíveis e os mais influentes nas mídias sociais. Ainda que alguns se sintam aliviados com essa constatação, ela não deveria trazer conforto. O sentimento de que a esfera pública é hoje um ambiente tóxico tomado por um diálogo de surdos não é uma ilusão criada pelas mídias sociais, que distorceriam uma realidade geral mais nuançada. Esses poucos que estão muito polarizados são aqueles que, por seu poder e influência, estruturam e organizam o debate público, tanto nas novas como nas velhas mídias. São eles também que, no final, orientam e informam os menos engajados, para o bem ou para o mal. As mídias sociais não parecem ser a causa da polarização política, nem nos EUA nem no Brasil. Mas o problema existe e não é uma miragem”.

## II. FAKE NEWS E DEEP FAKES: A REALIDADE EXISTE?

■ As chamadas fake news, notícias falsas criadas com o objetivo de desinformar, estão atingindo os usuários com precisão maior do que se espera. O tipo de conteúdo enviado também pode levar em conta o perfil pessoal de quem vai ler a notícia, a fim de causar impacto mais direto, o que é propiciado pela técnica de *microtargeting*.

É cada vez mais evidente o fato de que dados produzidos por usuários na internet estão sendo coletados de alguma forma por terceiros. Não apenas dados pessoais, como também o que leem, pesquisam e, principalmente, seus hábitos de consumo. Ao mesmo tempo, a internet possibilita a captação maciça desses dados e seu processamento em grande escala. Esse grande volume de dados – estruturados, semiestruturados ou não estruturados (Lane, 2014) – forma o *big data*, tecnologia que possibilita conhecer cada vez mais os indivíduos, podendo até mesmo identificá-los pessoalmente por meio da observação de seus hábitos, preferências e desejos.

A riqueza dessas informações é tamanha que se torna inevitável questionar como os usuários permitem tal coleta ao consentirem, por exemplo, com os termos de uso de sites e aplicativos. Ocorre, em primeiro lugar, que os termos de uso costumam ser extremamente técnicos e ininteligíveis para a população em geral, o que faz com que o consentimento dado não seja completamente consciente. Em segundo lugar, a atuação das empresas em si nem sempre é feita de forma transparente, ou seja, frequentemente a finalidade real destinada aos dados é escondida dos usuários<sup>8</sup>.

Com isso e com a quantidade crescente de dados produzida diariamente, torna-se preocupante o manejo dessas informações por terceiros. Isto porque o *big data* vai muito além de um emaranhado de dados: ele é essencialmente relacional<sup>9</sup>. À medida que os indivíduos não possuem controle de seus próprios dados pessoais, pode-se dizer que estes pertencem àqueles que os coletam, criando uma prejudicial relação vertical.

Tal tecnologia abre uma oportunidade que não passou despercebida para o mercado. Enxergou-se nesse volume de dados a possibilidade de personalização automática de conteúdo nas plataformas digitais, inclusive direcionando essa filtragem por meio de publicidade direcionada, feita através do rastrea-

8 Sobre os termos de uso na internet, cf. Magrani et. al, 2016.

9 Em sentido similar, cf. Santos, 2017.

mento de cookies e processos de *retargeting*, ou mídia programática (*behavioral retargeting*).

Empresas observam os insumos gerados por esses dados para orientar sua política de mercado, a fim de alcançar os desejos e hábitos dos consumidores, por meio de técnicas como *tracking*, *profiling* e *targeting*. Isto é feito de acordo com as tendências de comportamento analisadas, o que leva a um direcionamento, portanto, das escolhas de mercado mediante a criação de alvos. Observamos, hoje, a predominância nas esferas conectadas dos lucrativos modelos de negócio baseados em filtragem algorítmica com a finalidade de direcionar a venda de produtos e serviços de forma otimizada a *e*-consumidores.

A técnica de *microtargeting* é uma estratégia digital de criação de público-alvo por meio da coleta de dados desse público, para que a empresa possa conhecer minuciosamente o perfil em questão. A estratégia é feita em cima de um banco de dados montado com informações como idade, gênero, hobbies, comportamento, entre outros. Em princípio, o *microtargeting* era usado no marketing publicitário para o aprimoramento de produtos e serviços. Agora, fala-se de marketing político, uma vez que auxilia os candidatos a definirem um nicho de eleitores específicos, mapeando possíveis apoiadores.

Uma das vantagens do *microtargeting* é permitir antecipar resultados que possam ser atingidos ao final do projeto publicitário ou político, proporcionando economia de tempo e dinheiro por parte dos agentes, uma vez que seu foco será qualitativo em cima do que os alvos de fato querem ou precisam, dispensando tentativas aleatórias. Essas práticas correntes, portanto, pautam-se pela utilização em grande medida dos dados dos usuários através do *big data* que, além de fazerem uso desonesto de informações pessoais, também podem gerar consequências políticas, como o agravamento do efeito denominado “*filter bubble*”, nocivo ao papel democrático da internet como esfera pública, e a potencialização das notícias falsas<sup>10</sup>.

Sobre esse contexto político-democrático, alguns exemplos podem ajudar a entender como o *microtargeting* é usado para potencializar notícias falsas. O caso paradigmático é o das eleições de 2016 nos Estados Unidos, duramente impactada pelas fake news. Afirma-se que os boatos assumiram majoritariamente um teor negativo a respeito da candidata democrata Hillary Clinton, em contrapartida a um encorajamento da conduta do republicano Donald Trump. Fato é

---

10 Sobre a relação entre fake news e eleições, recomenda-se a leitura da carta aberta defendida pelo grupo Coalizão de Direitos na Rede (2017), trazendo *guidelines* sobre o assunto.



que, em 2016, 33 das 50 notícias falsas mais espalhadas no Facebook tratavam do contexto político vivido nos Estados Unidos (Existem 7 tipos, 2018; Silverman, 2016).

Algumas notícias falsas tiveram tamanha repercussão que rodaram o mundo, como, por exemplo, a de que o Papa Francisco – e, por consequência, a Igreja Católica Romana – apoiou a candidatura de Donald Trump, o que lhe renderia um apoio ainda maior das camadas conservadoras da sociedade americana. O boato foi desmentido apenas quando o porta-voz do Vaticano fez um pronunciamento público dizendo que o Papa nunca manifestou tal apoio e, tampouco, tem a intenção de assumir posições políticas.

Países como a Rússia também influenciaram no processo eleitoral americano. Dentre os boatos espalhados, um exército de “trolls russos” publicou notícias de que Hillary Clinton estaria envolvida com práticas de ritual satânico. Uma das linhas narrativas propagadas alegava, em cima de supostos e-mails vazados entre Hillary e seu gerente de campanha, John Podesta, que estes participavam de rituais com uma sacerdotisa que adorava o demônio. Tratava-se, no entanto, de uma performance da artista Marina Abramovic sobre *spirit cooking*, que foi questionada em um dos e-mails invadidos da campanha (Lee, 2016). Mais tarde, a inteligência dos EUA descobriu que os e-mails foram hackeados em uma operação orquestrada pelo Kremlin (How Russia-linked, 2017).

Porém, um caso recente que se tornou emblemático e, de fato, despertou atenção sobre como o *microtargeting* pode ser utilizado para disseminar notícias falsas, é o da empresa *Cambridge Analytica*, apontada como um dos principais vetores de viralização de fake news, assim como pela vitória de Donald Trump nas eleições.

O caso se inicia com a criação de um aplicativo que operava no Facebook, o *thisisyourdigitallife*, criado pelo acadêmico da Cambridge Analytica, Dr. Aleksandr Kogan, atuante da Universidade de Cambridge, com o objetivo de desenvolver pesquisas acadêmicas. Para isto, o *app* coletava informações privadas dos perfis de 270 mil usuários, com seu consentimento, o que até então era permitido e estava de acordo com os termos de uso do Facebook. Ocorre que, em 2015, a rede social em questão foi informada de que a Cambridge Analytica havia compartilhado os dados coletados com um terceiro, a empresa Eunoia Technologies, que visava fins comerciais, em desacordo com os termos de uso da plataforma. Dessa forma, o Facebook exigiu que as informações dadas a terceiros fossem destruídas, porém, posteriormente, descobriu-se que a Cambridge Analytica e outras empresas não eliminaram as informações, razão pela qual esta-

riam suspensos de operar na plataforma a partir daquele momento. Nesse ponto, porém, os dados de cerca de 50 milhões de usuários do Facebook já haviam sido utilizados.

O escândalo só veio a público em março de 2018, quando Christopher Wylie, que trabalhava para obter dados de usuários no Facebook e repassar para a empresa Cambridge Analytica (contratada internacionalmente por diversos políticos em tempos eleitorais), concedeu depoimentos à imprensa, revelando que os perfis foram colhidos para fins de manipulação política na esfera pública conectada (Cadwalladr, Graham-Harrison, 2018).

Esse caso desperta atenção para alguns fatores importantes. O *app thisisyour-digitallife* funcionava como um teste de personalidade que recompensava, ainda, financeiramente aqueles que aceitassem participar. Isso representa um apelo forte ao usuário de redes sociais, que tende a querer saciar a curiosidade do resultado desses testes, os quais se tornaram tão comuns, ainda mais mediante a possibilidade de auferir algum lucro com isso. De forma mascarada, portanto, a empresa conseguiu coletar grande quantidade de dados, de forma consentida para o uso de uma finalidade distinta.

A finalidade oculta, descobriu-se posteriormente, era de coletar dados para traçar perfis de eleitores, a fim de utilizá-los para marketing eleitoral. Isso nada mais é do que uma estratégia de *microtargeting*, fazendo uso da tecnologia do *big data* para alcançar um material mais refinado, capaz de produzir resultados ainda mais precisos.

A empresa gastou cerca de US\$ 1 milhão na coleta de dados para enviar mensagens direcionadas a eleitores específicos, manipulando sua opinião política através de um algoritmo que conseguia analisar os perfis individuais e determinar traços de personalidade ligados ao comportamento *online* do eleitor, bem como seus sentimentos e medos, direcionado o conteúdo de manipulação sociopolítica com base nesses fatores. Assim, uma gama preciosa de dados coletados pela Cambridge Analytica foi vendida para partidos políticos para, a partir de uma análise, produzir fake news capazes de atingir o eleitor naquilo que lhe é mais caro. Ou seja, corroborar ou atacar seus posicionamentos mais arraigados, com o propósito de dissuadi-los, com a certeza do êxito.

Com isto, é preciso estar claro que, em última instância, o *big data* é o indivíduo em toda sua complexidade e, portanto, deve-se ter uma consciência crítica e pensar possibilidades de retomar o controle sobre os dados pessoais (Magrani, Oliveira, 2018). Faz-se necessário dirigir-se a formas judiciais e extrajudiciais de

proteção de dados, da responsabilização de empresas que realizam tal atividade. E, principalmente, edificar um uso consciente das plataformas nos usuários, para que não abram mão tão facilmente de suas informações em falsas trocas de benefício, que acabam por voltar-se contra estes de forma tão penosa para a sociedade e para a democracia de forma geral.

De forma similar à campanha de Donald Trump em 2016, a campanha de Jair Bolsonaro em 2018 no Brasil utilizou inúmeras fake news para promoção do candidato. A estratégia se tornou pública quando a imprensa divulgou a existência de contratos do candidato com empresas privadas totalizando cerca de 12 milhões de reais através dos quais as empresas compravam pacotes de disparos de mensagens contra o partido oposito (PT) no WhatsApp, o que incluía a divulgação de notícias falsas (Redação Folha de São Paulo, 2018). O candidato também contou com a participação de grupos de pessoas voluntárias que organizavam a criação e divulgação de fake news (Simões, 2018). As notícias falsas com maior repercussão durante as eleições diziam respeito ao “Kit Gay” e à fraude nas urnas, e outras notícias que circularam menos tratavam da acusação de que Fernando Haddad (PT) seria pedófilo e de que Jair Bolsonaro (PSL) teria vontade de alterar a padroeira do Brasil (FGV DAPP, 2018c).

Outra forma a que se vem recorrendo para conquistar o eleitorado é a *deep fake*. As tecnologias já permitem a gravação de áudios com imitação praticamente idêntica à voz das pessoas e a edição de vídeos em que o rosto de um indivíduo que jamais estivera na situação veiculada aparece como partícipe. Se no cenário cotidiano de pessoas não públicas isso já se mostra extremamente prejudicial à honra e à imagem, esse risco cresce exponencialmente quando falamos de pessoas públicas. Áudios e vídeos editados podem ser utilizados, por exemplo, para difamar a imagem de determinado candidato a um cargo eleitoral.

Um caso recente ilustra essa possibilidade. No dia 23 de outubro de 2018, passou a circular na internet um vídeo em que, supostamente, o candidato a governador do estado de São Paulo, João Doria (PSDB), aparecia em cenas íntimas com mulheres. A cinco dias do segundo turno das eleições, a divulgação de um vídeo nesse sentido mostra-se extremamente prejudicial para a imagem do candidato, sobretudo quando se considera que Doria é defensor da família tradicional. O então candidato entrou com pedido de investigação na Justiça Eleitoral. Inicialmente, as investigações em relação ao vídeo apontaram que se trataria de montagem ou simulação: laudo pericial afirmou que o rosto do candidato teria sido inserido indevidamente no vídeo, colocando-o em situação da qual não par-

ticipara (Quintella, 2018). Posteriormente, novo laudo pontuou a veracidade do vídeo (Redação Pragmatismo, 2018)<sup>11</sup>.

Trata-se de caso paradigmático de *deep fakes*. Para além disso, a própria realidade é colocada em xeque, e não se sabe mais o que é verdadeiro ou falso. Isso cria uma confusão mental no eleitorado, que passa a acreditar em um dos lados sem qualquer fundamentação sólida. Tudo sendo questionável, o desejo humano por uma resposta se agarra a qualquer indício de veracidade – seja esse indício amparado em algum fato comprovado ou apenas no autoengano (Giannetti, 2005).

Após a divulgação do vídeo, as pesquisas de intenção de voto mostraram alguma variação nos pontos percentuais de cada candidato. De acordo com a pesquisa do Datafolha, em 25 de outubro de 2018, Doria contava com 52% de intenção de votos, ao passo que no dia 27 do mesmo mês havia caído para 49% (Fujita, 2018). Considerando a intensidade das campanhas nos dias imediatamente anteriores às eleições e a profusão de informações que são divulgadas, não podemos afirmar que o vídeo foi diretamente responsável por essa queda. Além disso, o primeiro laudo pericial divulgado indicava que o vídeo se trataria de uma montagem ou simulação, o que pode ter causado mais dúvidas no eleitor. Fato é que a divulgação dessa *deep fake*, acompanhada dos laudos periciais que não indicaram uma solução única, não foram suficientes para impedir a vitória do candidato, que venceu com 51,77% dos votos válidos. Note, contudo, que podemos afirmar que o vídeo se tratou de importante fator a ser enfrentado nos momentos finais da campanha. No contexto atual, os cidadãos têm consciência de que há divulgação indiscriminada de fake news, de modo que podem considerar, sem qualquer prova em nenhum dos sentidos, que a divulgação do vídeo se tratou de mera artimanha da oposição para descredibilizar o adversário. Assim, ignoram se o vídeo se tratou de fato verdadeiro ou falso e se prendem às crenças já formadas – as quais, muitas vezes, se baseiam em fake news.

Nesse contexto, surgem projetos de lei<sup>12</sup> que buscam criminalizar a divulgação de fatos falsos durante o ano eleitoral, como os Projetos de Lei da Câmara nº 9973/2018, 10292/2018, 9931/2018 e 9532/2018. O Projeto de Lei do Senado nº 246/2018 é mais amplo e busca inserir no Marco Civil da Internet “medidas de combate à divulgação de conteúdos falsos (fake news) ou ofensivos em aplicações de internet”. Além disso, crescem os grupos destinados a realizar *fact checking*.

---

11 No momento em que este estudo foi finalizado, o caso ainda não havia sido solucionado e apenas os laudos mencionados tinham sido divulgados.

12 Uma enumeração exaustiva e apresentação pormenorizada de todos os projetos de lei sobre o assunto demandaria estudo próprio e desbordaria dos limites deste artigo.

Mas em um cenário em que tudo é questionável, quem irá checar a veracidade da checagem sobre a realidade? A profusão de informações verdadeiras e falsas poderia levar a um “infocalipse”, como alerta Aviv Ovadya (2018). É por isso que afirmamos acima ser necessário garantir um patamar mínimo de consenso sobre a realidade e um respeito a princípios éticos básicos.

Os impactos dessa manipulação da esfera pública vão muito além das eleições. Em longo prazo, pode ser que a elaboração de políticas públicas se dê com base numa vontade popular forjada, gerando gastos estatais que não vão ao encontro da real necessidade dos cidadãos. Além disso, a própria legitimação constante da atuação dos políticos pode ser forjada, ainda que políticas públicas desinteressantes sejam colocadas em prática. No item a seguir, tecemos algumas considerações sobre o *filter bubble* e seu impacto na formação de convicção dos indivíduos e na configuração da esfera pública.

### III. A ESFERA PÚBLICA FORJADA POR ALGORITMOS E A CONVICÇÃO PESSOAL NA ERA DO *FILTER BUBBLE*<sup>13</sup>

■ O *Filter Bubble* (ou filtros-bolha) (Pariser, 2011) pode ser definida como um conjunto de dados gerado por todos os mecanismos algorítmicos, utilizados para se fazer uma edição invisível voltada à customização da navegação *online*. 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. Forma-se então, a partir das características de navegação de cada pessoa, um universo particular *online*, condicionando sua navegação. Isto se dá por meio do rastreamento de várias informações, dentre elas, a localização do usuário e o registro dos *cookies* (Wu, 2011)<sup>14</sup>.

Com essas técnicas que geram os filtros-bolha, a internet estaria se transformando em um espaço no qual é mostrado o que se acha que é de nosso interesse. Assim, quase sempre nos é ocultado aquilo que de fato desejamos ou eventualmente precisamos ver. Assim, pode-se dizer que a *filter bubble* é paternalista e prejudicial para o debate e a formação de consenso na esfera pública conectada, sendo possível até questionar sua constitucionalidade, já que pode implicar restri-

13 Algumas das considerações feitas nesse capítulo foram exploradas em Magrani, 2018a e 2014.

14 *Cookies* são, em poucas palavras, dados de acesso que consistem nas “pegadas digitais” deixadas ao se transitar e se manifestar pelos ambientes *online*.

ções a direitos fundamentais, como acesso à informação, liberdade de expressão, bem como à própria autonomia dos indivíduos<sup>15</sup>.

A filtragem surgiu como uma necessidade e é, muitas vezes, considerada bem-vinda, gerando um comodismo muito grande ao usuário, que encontra de forma rápida e eficaz, em grande parte das vezes, a informação ou qualquer outro conteúdo que deseja acessar. Este é o modelo de negócios do *Netflix*, por exemplo, que permite que o usuário tenha à sua disposição um acervo de filmes baseado unicamente no seu perfil através da sugestão de títulos e filtros personalizados, com intuito de melhorar a sua experiência.

No entanto, para além da conveniência, o problema reside na forma e no excesso da filtragem, tanto por parte das empresas quanto dos próprios indivíduos que, sem ter consciência, se limitam e se afastam de pontos de vista divergentes dos seus, empobrecendo, assim, o valor do debate na esfera pública virtual. Por isso, os filtros-bolha limitam os usuários ao que desejam (ou desejariam) segundo, na maior parte das vezes, uma predição algorítmica (Morozov, 2013). Isso dificulta o acesso às informações que deveriam ou precisariam ser vistas para o enriquecimento do debate democrático.

Ademais, em outra perspectiva, o usuário de internet, ao navegar pelos sites mais conhecidos, é alvo hoje de uma torrente de publicidade direcionada que denota por si só o interesse comercial por trás deste mecanismo de filtragem e personalização.

A internet é plástica e mutável e o fato de nos tornarmos involuntariamente reféns dos algoritmos que nos inserem dentro destas bolhas tem sido encarado como uma das mudanças mais drásticas, mas sutis, por serem muitas vezes imperceptíveis. A premissa do *filter bubble* é que o usuário não decide deliberadamente o que aparece para ele dentro da bolha, nem tem acesso ao que fica de fora.

A curadoria de informação realizada pela mídia tradicional, nos meios *offline* inclusive, já concretiza a ideia de filtragem de conteúdo selecionando e segregando uma série de informações. Habermas, assim como outros teóricos da Escola de Frankfurt, como Adorno e Horkheimer (Wiggershaus, 1995), já atentava para a força da mídia tradicional e seu impacto para a democracia moderna (Habermas, 2003, p. 99). No entanto, muitas vezes as plataformas de internet não possuem transparência suficiente no recorte informacional e algorítmico que realizam, dando uma falsa ideia ao consumidor de que as informações possuem um fluxo

---

15 Uma afirmação peremptória nesse sentido demandaria estudo mais aprofundado, de modo que a abordagem específica desse ponto desbordaria dos limites desse estudo.

neutro e livre. Além disso, a filtragem por algoritmos que se vê nos ambientes *online* permite um grau de personalização e direcionamento em uma escala muito maior (Magrani, 2014), o que tende a se intensificar com o advento da Internet das Coisas (cf. Magrani, 2018b), tendo em vista que com cada vez mais dispositivos inteligentes conectados ao nosso redor, teremos ainda mais dados pessoais sendo recolhidos, armazenados e tratados.

Diante do exposto até aqui, percebe-se que a ideia de que a infraestrutura da internet como esfera pública tem o potencial de permitir que as discussões possuam força suficiente para chegar a diferentes segmentos e a grupos de interesses diversos, replicando-se pelas várias redes de pessoas que compõem a sociedade, talvez seja uma realidade cada vez mais distanciada. Isso se deve ao fato de que as expressões ficam muitas vezes restritas a uma mesma rede de pessoas com interesses comuns e com canais de comunicação facilmente manipuláveis pelos detentores das plataformas. A consequência disto é a intensificação da fragmentação comunicacional e a polarização do debate público<sup>16</sup>.

Em uma visão habermasiana de legitimação do sistema político-democrático, este cenário é condenável, uma vez que o fluxo comunicacional minimamente livre deve ser preservado no espaço público, permitindo que todos os possíveis atingidos tenham voz e participem de forma cada vez mais direta nas decisões, sejam elas pertinentes ao seu contexto privado ou politicamente na esfera pública. Um exemplo paradigmático disso é o caso da Cambridge Analytica, descrito no item precedente.

Com o ganho de maior sofisticação e autonomia das tecnologias, nossa interação com esses agentes ficará cada vez mais simbiótica e complexa, trazendo à tona, ainda, uma maior capacidade de manipulação do nosso pensamento e comportamento.

Devemos somar a isso, como algo negativo, o fato de que, muitas vezes, não conhecemos como os algoritmos dos objetos inteligentes que utilizamos e dos espaços virtuais onde interagimos funcionam<sup>17</sup>. Cada vez mais esses novos agentes não-humanos produzem efeitos em nossas ações ou mesmo tomam decisões

---

16 Como nota Cass Sunstein (2001 e 2009), os filtros-bolha seriam um sério risco para o potencial da esfera pública conectada devido à falta de contato com opiniões dissidentes e à polarização dos discursos levando a radicalismos. Este seria um problema com tendências não à sua resolução, mas ao seu agravamento, a partir da sofisticação dos algoritmos de personalização de conteúdo.

17 Frank Pasquale (2015) faz uma crítica a essa situação, tratando os algoritmos de hoje como caixas pretas e jogando luz sobre os efeitos disso em uma sociedade guiada em diversas áreas por dados e decisões algorítmicas.

importantes em nosso lugar através de customização da informação que nos é oferecida<sup>18</sup>.

De forma geral, a tomada de decisões e a interação democrática comunicativa hoje estão passando por uma transformação profunda, pois sofrem a intermediação e o agenciamento de agentes não-humanos, como robôs ou algoritmos dotados de algum grau de inteligência artificial. Esses elementos estão influenciando nossa interação e nosso discurso com capacidade de produzir efeitos materiais de cunho político-democrático significativos, por isso devem ser melhor compreendidos para fins de regulação.

Nas discussões políticas, os robôs têm sido usados por todo o espectro partidário não apenas para conquistar seguidores, mas também para conduzir ataques a opositores e forjar discussões. Eles manipulam debates, criam e disseminam notícias falsas e influenciam a opinião pública ao postar e replicar mensagens em larga escala. Muitos *bots*<sup>19</sup> (robôs) têm replicado *hashtags* no Twitter<sup>20</sup> e no

---

18 Em 2017, em Wisconsin no EUA, um juiz concedeu uma pena de prisão de seis anos levando em consideração não apenas o registro criminal do réu, mas também sua pontuação na escala COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*), que é uma ferramenta algorítmica que visa prever o risco de reincidência de um indivíduo. A pontuação sugeriu que o réu tinha um alto risco de cometer outro crime; assim, sua sentença foi de seis anos. O réu apelou da decisão, com o argumento de que o uso pelo juiz do algoritmo preditivo em sua decisão de sentença violou o devido processo e se pauta pela opacidade dos algoritmos. Cf. Liptak, 2017. O caso foi para a Suprema Corte dos Estados Unidos, a qual negou o *writ of certiorari*, recusando-se a analisar o caso.

19 O termo *Bot*, diminutivo de *Robot* (ou *Internet bot* ou *web robot*), é uma aplicação de software que tem o objetivo de oferecer um serviço automatizado para realizar tarefas em geral predefinidas. Eles imitam comportamentos humanos e vêm sendo utilizados na política e nas eleições para influenciar opinião em redes digitais, como em plataformas de redes sociais, mensagens instantâneas ou sites de notícias. Uma conceituação do termo pode ser encontrada em Velasco, Domingos, 2017.

20 De acordo com o projeto PEGABOT, do Instituto de Tecnologia e Sociedade do Rio de Janeiro (ITS Rio) e do Instituto Equidade & Tecnologia, “[u]m Twitter Bot é uma conta controlada por um algoritmo ou script, normalmente utilizadas para realizar tarefas repetitivas, por exemplo, retweetar conteúdo contendo palavras-chave particulares, responder a novos seguidores e enviar mensagens diretas a novos seguidores. Twitter Bots mais complexos podem participar de conversas online e, em alguns casos, tem um comportamento muito parecido ao comportamento humano. As contas bot compõem entre 9 e 15% de todas as contas ativas no Twitter, mas estudos mais aprofundados indicam que este percentual pode ser ainda maior devido a dificuldade de identificar os bots complexos. Os bots do Twitter geralmente não são criados com intenção maliciosa; eles são freqüentemente usados para melhorar a interação on-line ou na prestação de serviços por empresas, governos e outras organizações, por isso é importante separar os bons bots dos maus bots”. Disponível em: <<https://pegabot.com.br>>. Acesso em: 27 out. 2018.



Facebook<sup>21</sup> que ganham destaque com a massificação de postagens automatizadas de forma a sufocar debates espontâneos sobre um determinado tema.

A princípio, as contas automatizadas podem até contribuir positivamente em alguns aspectos da vida nas redes sociais. Os *chatbots* (chats operados por robôs) (Instituto Tecnologia & Equidade, 2017), por exemplo, agilizam o atendimento a clientes de empresas e, em alguns casos, até auxiliam consumidores a processarem seus pedidos e obterem mais informações. Porém, um número crescente de robôs atua com fins maliciosos na esfera pública. Os *social bots* (robôs sociais) são contas controladas por softwares, que geram conteúdo artificialmente e estabelecem interações com não robôs. Eles buscam imitar o comportamento humano e se passar como tal de maneira a interferir em debates legítimos e voluntários e criar discussões forjadas (Ruediger, 2017).

O crescimento da ação protagonizada por robôs representa, portanto, uma ameaça real para o debate público, representando riscos, no limite, à própria democracia, ao manipular o processo de formação de consensos na esfera pública e de seleção de representantes e agendas de governo<sup>22</sup>. Não por outro motivo, há projetos de lei tramitando no Brasil em âmbito federal para proibir a utilização e a contratação de *bots* para fins eleitorais, como o Projeto de Lei do Senado nº 413/2017, que busca criminalizar “a oferta, a contratação ou a utilização de ferramenta automatizada que simule ou possa ser confundida com pessoa natural para gerar mensagens ou outras interações, pela internet ou por outras redes de

---

21 Os robôs têm maior facilidade de propagação no Twitter do que no Facebook por uma série de motivos. Uma explicação sobre o assunto pode ser conferida em Ruediger, 2017.

22 Bots são responsáveis por mais de 50% do tráfego da internet ao redor do mundo. Alguns bots têm como propósito, por exemplo, exigir prestação de contas de políticos, viralizar causas para a igualdade de gênero ou ajudar a organizar as (muitas) tarefas diárias de seus usuários. Já outros bots têm como objetivo espalhar mentiras para influenciar conversas na esfera pública, um fenômeno que desde 2014 vem ganhando escala global. Esses bots estão por aí e quase ninguém sabe como eles funcionam, quem os desenvolve e por quem são financiados. Para ilustrar essa questão, uma pesquisa recente demonstrou que a repercussão do cancelamento do evento do Queermuseu, exaustivamente comentada na imprensa nacional, foi insuflada por robôs na internet. Dos mais de 700 mil tweets analisados, 8,69% foram disparados por bots, prejudicando a discussão pública. “Embora a decisão pelo cancelamento da exposição tenha levado em conta outros fatores, é possível dizer que a ação dos bots impactou na forma com que o debate foi conduzido, e suas consequências práticas. (...) O uso dos bots provoca um ambiente de polarização, uma vez que a internet tem um aumento no fluxo de mensagens com o mesmo teor. Neste cenário, assegura o pesquisador, fica difícil surgir um debate espontâneo, com ideias discordantes e moderadas. ‘Esse tipo de ação dificulta o surgimento de posições mais moderadas. A busca de um consenso fica prejudicada porque os robôs conseguem sequestrar parte do debate’”. Cf. G1 RS, 2017.

comunicação, com o objetivo de influenciar o debate político ou de interferir no processo eleitoral”.

Corroborando a tese de risco à democracia, a Diretoria de Análise de Políticas Públicas (DAPP) da FGV identificou interferências ilegítimas no debate *online* através do uso de *bots* nas eleições de 2018 (Ruediger, 2017) e de 2014 (FGV DAPP, 2018b) e em debates públicos de forma geral (Ruediger, 2018). Contas programadas para postagens massivas se converteram em uma ferramenta para a manipulação de debates nas redes sociais. Aqui, é importante ressaltar que a mídia tradicional, sobretudo a televisão, vem sofrendo um constante processo de desgaste e de descrédito por parte dos cidadãos. Nesse contexto, os indivíduos passam cada vez mais a utilizar a internet para se informar e a confiança nos dados obtidos através do computador se mostra superior em relação aos demais meios de comunicação, como jornais, rádios e televisão (Brasil, 2016). Contudo, o cenário *online* está permeado por *bots* e algoritmos que forjam o debate e alteram a prioridade dos temas. Durante a corrida eleitoral de 2018, as contas automatizadas chegaram a ser responsáveis por 12,9% das interações no Twitter (FGV DAPP, 2018a). Em 2014, primeira eleição presidencial em que os robôs tiveram atuação mais substancial, a interferência foi similar. Os *bots* responderam por mais de 10% das interações no Twitter. Já durante o processo de Impeachment da Ex-Presidenta Dilma Rousseff, os robôs foram responsáveis por 20% do debate entre apoiadores de Dilma. No segundo turno das eleições de 2014, 20% das interações a favor de Aécio Neves foram geradas por robôs (Ruediger, 2017).

Com este tipo de manipulação, os robôs criam a falsa sensação de amplo apoio político a certa proposta, ideia ou figura pública, modificam o rumo de políticas públicas, interferem no mercado de ações, disseminam rumores, notícias falsas e teorias conspiratórias, geram desinformação e poluição de conteúdo, além de atrair usuários para links maliciosos que roubam dados pessoais, entre outros riscos<sup>23</sup>. Note, contudo, que dizer que esses *bots* atuam em favor de determinada agenda não significa que “dominem completamente a rede nem que a percepção final da maior parte das pessoas será resultante direta da influência desses dispositivos” (Ruediger, 2017). O que buscamos destacar são os perigos já concretizados através do uso de robôs e os riscos potenciais que se mostram cada vez mais próximos e temerários.

Ao interferir em debates em desenvolvimento nas redes sociais, robôs estão atingindo diretamente os processos políticos e democráticos através da influência

---

23 Sobre a existência hoje de um “exército” de perfis falsos, cf. Gagnani, 2017.

da opinião pública. Suas ações podem, por exemplo, produzir uma opinião artificial, ou dimensão irreal de determinada opinião ou figura pública, ao compartilhar versões de determinado tema, que se espalham na rede como se houvesse, dentre a parcela da sociedade ali representada, uma opinião muito forte sobre determinado assunto (Córdova, Doneda, 2017).

O estudo do uso de robôs já demonstra de forma clara o potencial danoso dessa prática para a disputa política e o debate público<sup>24</sup>. Uma das conclusões mais evidentes nesse sentido é a concentração dessas ações em polos políticos localizados no extremo do espectro político, promovendo artificialmente uma radicalização do debate nos filtros-bolha e, conseqüentemente, minando possíveis pontes de diálogo entre os diferentes campos políticos constituídos. Assim, a atuação de robôs não apenas dissemina notícias falsas, que podem ter efeitos nocivos para a sociedade, mas também busca ativamente impedir que os usuários se informem de maneira adequada.

Outra estratégia comum dos perfis automatizados é o compartilhamento de links maliciosos, que tem como fim o roubo de dados ou informações pessoais. Essas informações – como fotos de perfil – podem ser usadas para a criação de novos perfis robôs que tenham características que os auxiliem a iniciar conexões nas redes com usuários reais. Uma ação comum, que costuma gerar suspeita sobre a atuação de robôs, é a marcação por parte de um usuário desconhecido.

Este tipo de atuação sugere que as redes sociais, usadas por tantas pessoas para fins de informação, podem estar, na verdade e de forma paradoxal, contribuindo para uma sociedade menos informada, manipulando o debate público. Somados, esses riscos e outros representados pela ação de artefatos *não-humanos* (como *bots*), são mais do que o suficiente para jogar luz sobre uma ameaça real à quali-

24 Segundo a pesquisa (Ruediger, 2017): “A detecção através de aprendizado de máquinas ocorre com a codificação de padrões de comportamento a partir da coleta de metadados. Desta forma, o sistema é capaz de identificar automaticamente humanos e robôs com base no padrão comportamental do perfil. Os metadados dos usuários são considerados um dos aspectos mais previsíveis para diferenciar humanos e robôs e podem contribuir para uma melhor compreensão do funcionamento de robôs mais sofisticados. Identificar esses robôs ou contas hackeadas, no entanto, é difícil para estes sistemas. Além disso, a evolução constante dos robôs faz com que o sistema, construído a partir de uma base de dados estática, se torne menos preciso ao longo do tempo. No entanto, ele permite processar um grande número de correlações e padrões complexos, além de analisar um grande número de contas. Os mecanismos mais eficientes de identificação combinam diferentes aspectos dessas abordagens, explorando múltiplas dimensões do comportamento do perfil, como atividade e padrão de horário. Estes sistemas levam em conta, por exemplo, que usuários reais passam mais tempo na rede trocando mensagens e visitando o conteúdo de outros usuários, como fotos e vídeos, enquanto contas robôs passam o tempo pesquisando perfis e enviando solicitações de amizade”.

dade do debate na esfera pública<sup>25</sup>, sobretudo porque os artefatos não-humanos vêm ganhando cada vez mais autonomia e imprevisibilidade comportamental<sup>26</sup>.

## CONSIDERAÇÕES FINAIS

■ Os acontecimentos recentes envolvendo as novas tecnologias abordadas neste estudo nos alertam para o fato de que o papel democrático da esfera pública conectada começa a esbarrar em riscos e obstáculos que podem reduzir consideravelmente seu potencial, além de não dever ser encarada entusiasticamente como a panaceia para a salvação da legitimidade do sistema político contemporâneo.

A influência hipertrófica da racionalidade econômica do mercado e burocrática do sistema político nas esferas do mundo da vida é encarada por Habermas como uma das principais patologias da modernidade, levando a perdas de liberdade e de sentido na sociedade.

Por isso, o inicial frenesi com o ideal de esferas virtuais democráticas e descolonização do mundo da vida propiciada pelos novos ambientes digitais tem perdido fôlego. Agora que os algoritmos e demais agentes não-humanos estão participando e influenciando os discursos na esfera pública, cabe a indagação: serão eles obrigados a agirem moralmente e de forma racional-dialógica na comunicação para não afetar negativamente a situação ideal de fala?

Muitas vezes não há uma consciência crítica sobre como os algoritmos que compõem as tecnologias funcionam e como podem nos oferecer informações personalizadas a partir dos nossos dados pessoais ou mesmo manipular nossa visão política. É importante termos em mente que, frequentemente, esse funciona-

---

25 Para Habermas (2003, p. 28-30), devemos atingir ao máximo as condições de fala ideal, isto é, conseguir criar um ambiente de deliberação democrática em que todos tenham voz. Diante de um cenário de crise de representatividade, a internet deve ser utilizada como ferramenta para que o cidadão exerça a sua cidadania de maneira ativa. Segundo Habermas, para a deliberação democrática ocorrer, há pelo menos quatro condições. Estas condições, caracterizadas de uma “situação ideal de fala”, estão atreladas basicamente à necessidade de se garantirem as melhores condições de deliberação e à preocupação com a forma como se organiza o processo de debate. São elas: (i) cada pessoa precisa estar hábil a expressar suas próprias ideias abertamente e criticar as dos outros; (ii) a associação dos conceitos de força e poder com status social precisa ser eliminada; (iii) argumentos baseados no apelo à tradição ou dogma precisam ser expostos; e, como consequência, a verdade é alcançada por meio da busca ao consenso.

26 Nesse sentido, é paradigmático o exemplo do robô Tay, *chatbot* com capacidade de *deep learning* criado em 2016 pela *Microsoft*. A experiência se mostrou desastrosa e o robô teve que ser desativado em menos de 24 horas depois do início de seu funcionamento: Tay passou a disseminar discursos de ódio contra minorias historicamente marginalizadas, tendo afirmado, por exemplo, que Hitler estava certo e que ela odiava judeus. Sobre o robô Tay, cf. Moreira, 2016.

mento atende a disputas políticas ou a modelos de negócio privados que visam maximizar lucro e não necessariamente concretizar direitos fundamentais como acesso à informação, expressão e cultura.

A teoria habermasiana fundamentada nos conceitos comunicacionais racionais e dialógicos de esfera pública e situação ideal de fala nos ajuda a observar o quanto estamos nos distanciando de um cenário positivo do ponto de vista da legitimidade democrática. Pela análise feita, podemos compreender a atual situação como uma colonização do mundo da vida reforçada por meio de agentes não-humanos (*bots*, algoritmos com inteligência artificial, entre outros) – e também por agentes humanos, na medida em que os indivíduos também compartilham e produzem fake news e *deep fakes* –, produzindo consequências nocivas agravadas pelos efeitos de filtro-bolha e de radicalização dos discursos. A regulação jurídica precisa estar atenta a esses efeitos, buscando corrigi-los.

No contexto eleitoral de 2018, as fake news, de modo particular, e as novas tecnologias, de modo geral, se mostraram um problema de difícil solução. Por um lado, controlar a divulgação e a circulação de notícias falsas após sua veiculação seria de eficácia extremamente duvidosa, por conta da rápida velocidade com que a informação é disseminada no contexto da sociedade da informação. De outro, fazer a análise prévia da veracidade das notícias divulgadas poderia implicar em formas de censura institucionalizada.

É preciso, portanto, pensar em formas institucionais de combate às fake news sem que um dos temores apontados acima se concretizem. Assim, é provável que regulações indiretas sejam mais eficientes para combater as fake news, como a proibição de inúmeras contas falsas e o estabelecimento de padrões éticos para o uso de algoritmos e da inteligência artificial.

Note, no entanto, que legislar sobre esses temas é extremamente complicado, já que estamos lidando com princípios básicos da democracia, como liberdade de expressão e direito de acesso à informação. Mas essa ainda parece ser a alternativa mais apropriada a curto e médio prazos. Há tecnologias que podem ser usadas em *smartphones* e computadores para averiguar a veracidade de algumas informações (Viana, Zanatta, 2018). Contudo, é uma tecnologia de valor elevado, que demanda infraestrutura e a reposição de dispositivos que já circulam atualmente. Ou seja, trata-se de medida de longo prazo e com muitas dificuldades a serem enfrentadas, como as atinentes à privacidade dos usuários da tecnologia.

Como podemos observar, a cada dia as novas tecnologias passam a exercer maior influência na vida dos cidadãos e na forma como eles olham para os fatos. Essa influência se expande cada vez mais para todas as áreas de nossas vidas e,

recentemente, tem atingido profundamente as eleições. Por mais que ainda não seja possível afirmar que a manipulação algorítmica, o uso de *bots*, a divulgação de fake news e de *deep fakes* são os grandes responsáveis pelos resultados eleitorais, podemos dizer que estamos caminhando para um cenário de processo eleitoral hackeável.

---

EDUARDO MAGRANI · Doutor e Mestre em Direito Constitucional pela Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio) e Senior Fellow na Universidade Humboldt de Berlim, no Alexander von Humboldt Institute for Internet and Society. Coordenador do Instituto de Tecnologia e Sociedade do Rio de Janeiro (ITS Rio). Pesquisador Associado na Law Schools Global League e membro do Global Network of Internet & Society Research Centers (NoC). Professor das disciplinas de Direito e Tecnologia e Propriedade Intelectual em universidades renomadas como FGV, IBMEC e PUC-Rio. Advogado atuante nos campos de Direitos Digitais, Direito Societário e Propriedade Intelectual. Autor de diversos livros e artigos na área de Direito e Tecnologia e Propriedade Intelectual. Dentre eles os livros *Democracia Conectada* (2014), *Digital Rights: Latin America and the Caribbean* (2017) e *A Internet das Coisas* (2018).

RENAN MEDEIROS DE OLIVEIRA · Mestrando em Direito Público e Bacharel em Direito pela Universidade do Estado do Rio de Janeiro (UERJ). Pós-graduando em Direito Público pela Pontifícia Universidade Católica de Minas Gerais (PUC Minas). Pesquisador no Centro de Justiça e Sociedade da Fundação Getúlio Vargas (CJUS/FGV) e na Clínica de Direitos Fundamentais da Faculdade de Direito da UERJ – Clínica UERJ Direitos. Pesquisador Permanente do Laboratório de Regulação Econômica da UERJ – UERJ.

## REFERÊNCIAS

BOHMAN, James; REHG, William. Jürgen Habermas. *Stanford Encyclopedia of Philosophy*, revised edition, 2014 [first published 2007]. Disponível em: <<https://plato.stanford.edu/entries/habermas/>>. Acesso em: 28 nov. 2017.

BRASIL. Presidência da República. Secretaria Especial de Comunicação Social. *Pesquisa brasileira de mídia 2016: hábitos de consumo de mídia pela população brasileira*. Brasília: Secom, 2016.

CADWALLADR, Carole; GRAHAM-HARRISON, Emma. Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. *The Guardian*, mar. 2018. Disponível em: <<https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election>>. Acesso em: 29 abr. 2017.

CALHOUN, Craig (Ed.). *Habermas and the Public Sphere*. The MIT Press, 1992.

COALIZAÇÃO DE DIREITOS NA REDE. *Fake news e eleições*. Carta aberta de representantes da sociedade civil da América Latina e do Caribe sobre as preocupações relativas ao discurso sobre “fake news” e eleições. Direitos na rede, dez. 2017. Disponível em: <<https://direitosnarede.org.br/p/carta-aberta-americalatinaecaribe-igf2017/>>. Acesso em: 29 out. 2017.

COHEN, Joshua. Deliberation and Democratic Legitimacy. In: BOHMAN, J. e REHG, W. (Ed). *Deliberative Democracy: essays on reason and politics*. Cambridge, MA: MIT Press, 1997.

CÓRDOVA, Yasodara e DONEDA, Danilo. Um lugar para os robôs (nas eleições). *JOTA*, 2017. Disponível em: <<https://www.jota.info/opiniao-e-analise/artigos/um-lugar-para-os-robos-nas-eleicoes-20112017>>. Acesso em: 09 mar. 2018.

EXISTEM 7 TIPOS de fake news. Você conhece todos? *Magic Web Design*, mar. 2018. Disponível em: <<https://www.magicwebdesign.com.br/blog/internet/existem-7-tipos-fake-news-voce-conhece-todos/>>. Acesso em: 29 out. 2018.

FGV DAPP. Debate influenciado por robôs chega a 10,4% no Twitter. *FGV DAPP*, out. 2018a. Disponível em: <<https://observa2018.com.br/posts/debate-influenciado-por-robos-volta-a-crescer-e-chega-a-104-das-discussoes-sobre-os-presidenciais-no-twitter/>>. Acesso em: 29 out. 2018.

\_\_\_\_\_. *Robôs, redes sociais e política no Brasil*. Análise de interferências de perfis automatizados nas eleições de 2014. Rio de Janeiro: FGV DAPP, 2018b.

\_\_\_\_\_. “Fraude nas urnas” e “kit gay” têm maior impacto que outras notícias falsas em Twitter, Facebook e Youtube. *FGV DAPP*, out. 2018. Disponível em: <<https://observa2018.com.br/posts/fraude-nas-urnas-e-kit-gay-tem-maior-impacto-que-outras-noticias-falsas-em-twitter-facebook-e-youtube/>>. Acesso em: 29 out. 2018.

FUJITA, Gabriela. SP: Datafolha mostra França com 51% e Doria, 49%; Ibope traz 50% para cada. *UOL*, out. 2018. Disponível em: <<https://noticias.uol.com.br/politica/eleicoes/2018/noticias/2018/10/27/datafolha-ibope-sp-doria-franca.htm>>. Acesso em: 29 out. 2018.

G1 RS. Pesquisa demonstra que repercussão do cancelamento do Queermuseu foi insuflada por robôs na internet. *G1*, fev. 2018. Disponível em: <<https://g1.globo.com/rs/rio-grande-do>>

sul/noticia/pesquisa-demonstra-que-repercussao-do-cancelamento-do-queermuseu-foi-insuflada-por-robos-na-internet.ghhtml>. Acesso em: 02 mar. 2017.

GIANNETTI, Eduardo. *Auto-engano*. São Paulo: Companhia das Letras, 2005.

GRAGNANI, Juliana. Exclusivo: investigação revela exército de perfis falsos usados para influenciar eleições no Brasil. *BBC News*, dez. 2017. Disponível em: <<https://www.bbc.com/portuguese/brasil-42172146>>. Acesso em: 14 mar. 2018.

HABERMAS, Jürgen. *Direito e Democracia: entre facticidade e validade*. v. II. 2. ed. Rio de Janeiro: Tempo Brasileiro, 2003.

\_\_\_\_\_. *The Theory of Communicative Action*. v. II. Beacon Press, 1987.

HOW RUSSIA-LINKED hackers stole the Democrats' emails and destabilised Hillary Clinton's campaign. *ABC News*, nov. 2017. Disponível em: <<https://www.abc.net.au/news/2017-11-04/how-russians-hacked-democrats-and-clinton-campaign-emails/9118834>>. Acesso em: 29 out. 2018.

INSTITUTO TECNOLOGIA & EQUIDADE. Especialistas explicam como o robô pode influenciar o debate nas redes. *Medium*, dez. 2017. Disponível em: <<https://medium.com/@tecnoequidade/especialistas-explicam-como-o-robô-pode-influenciar-o-debate-nas-redes-3a844f911849>>. Acesso em: 29 out. 2017.

LANE, Julia et al. (Eds.). *Privacy, big data and the public good: frameworks for engagement*. New York: Cambridge University Press, 2014.

LEE, Benjamin. Marina Abramović mention in Podesta emails sparks accusations of satanism. *The Guardian*, nov. 2016. Disponível em: <<https://www.theguardian.com/artanddesign/2016/nov/04/marina-abramovic-podesta-clinton-emails-satanism-accusations>>. Acesso em: 29 out. 2018.

LIPTAK, Adam. Sent to Prison by a Software Program's Secret Algorithms. *The New York Times*, may 2017. Disponível em: <<https://www.nytimes.com/2017/05/01/us/politics/sent-to-prison-by-a-software-programs-secret-algorithms.html?mtrref=www.google.com.br&gwh=B3F914oAAAB1DACDFCE11CBD55F4DB8F&gwt=pay>>. Acesso em: 29 out. 2017.

MAGRANI, Eduardo et. al. *Terms of Service and Human Rights: an analysis of online platform contracts*. Rio de Janeiro: Revan, 2016.

\_\_\_\_\_; OLIVEIRA, Renan Medeiros de. We are big data: new technologies and personal data management. *CyberLaw*, v. 5, p. 10-33, 2018.

\_\_\_\_\_. *A Internet das Coisas: Privacidade e Ética na Era da Hiperconectividade*. 2018. (Doutorado em Direito) – Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2018a.

\_\_\_\_\_. *A internet das coisas*. Rio de Janeiro: FGV Editora, 2018b.

\_\_\_\_\_. *Democracia conectada: a internet como ferramenta de engajamento político-democrático*. Curitiba: Juruá, 2014.

MOREIRA, Isabela. A Microsoft criou uma robô que interage nas redes sociais – e ela virou nazista. *Galileu*, mar. 2016. Disponível em: <<https://revistagalileu.globo.com/blogs/buzz/>>



noticia/2016/03/microsoft-criou-uma-robo-que-interage-nas-redes-sociais-e-ela-virou-nazista.html>. Acesso em: 29 out. 2018.

MOROZOV, Evgeny. *To Save Everything, Click Here: The Folly of Technological Solutionism*. New York: Ed Public Affairs, 2013.

ORTELLADO, Pedro. Polarização na internet não parece ser causada pelas “bolhas”. *Folha de São Paulo*, fev. 2018. Disponível em: <[https://www1.folha.uol.com.br/colunas/pablo-ortellado/2018/02/polarizacao-na-internet-nao-parece-ser-causada-pelas-bolhas.shtml?utm\\_source=facebook&utm\\_medium=social&utm\\_campaign=compfb](https://www1.folha.uol.com.br/colunas/pablo-ortellado/2018/02/polarizacao-na-internet-nao-parece-ser-causada-pelas-bolhas.shtml?utm_source=facebook&utm_medium=social&utm_campaign=compfb)>. Acesso em: 29 out. 2018.

OVADYA, Aviv. What’s Worse Than Fake News? The Distortion Of Reality Itself. *New perspectives Quaterly*, v. 35, n. 2, p. 43-45, 2018.

PARISER, Eli. *The Filter Bubble: What the internet Is Hiding from You*. London: Penguin Press, 2011.

PASQUALE, Frank. *The Black Box Society: The secret algorithms that control money and information*. Cambridge: Harvard University Press, 2015.

PRIVACIDADE NO FACEBOOK: o que aprender com a Cambridge Analytica. *Irisbh*, [s.d.]. Disponível em: <<http://irisbh.com.br/privacidade-no-facebook-cambridge-analytica/>>. Acesso em: 28 out. 2018.

QUINTELLA, Sérgio. Perícia revela laudo sobre vídeo íntimo atribuído a João Doria. *Veja São Paulo*, out. 2018. Disponível em: <<https://vejasp.abril.com.br/blog/poder-sp/pericia-aponta-montagem-em-video-intimo-atribuido-a-joao-doria/>>. Acesso em: 29 out. 2018.

RAMOS, Marcella. Pedofilia, estupro, incesto: dois boatos de difamação desmentidos por dia no 2º turno. *Piauí*, out. 2018. Disponível em: <<https://piaui.folha.uol.com.br/pedofilia-estupro-incesto-dois-boatos-de-difamacao-desmentidos-por-dia-no-2o-turno/>>. Acesso em: 29 out. 2018.

REDAÇÃO FOLHA DE SÃO PAULO. Empresários bancam campanha contra o PT pelo WhatsApp. *Folha de São Paulo*, out. 2018. Disponível em: <<https://www1.folha.uol.com.br/poder/2018/10/empresarios-bancam-campanha-contr-o-pt-pelo-whatsapp.shtml>>. Acesso em: 29 out. 2018.

REDAÇÃO PRAGMATISMO. Vídeo íntimo de João Doria é verdadeiro, aponta novo laudo. *Pragmatismo Político*, out. 2018. Disponível em: <<https://www.pragmatismopolitico.com.br/2018/10/video-intimo-joao-doria-verdadeiro-pericia.html>>. Acesso em: 29 out. 2018.

RUEDIGER, Marco Aurélio (Coord.). Robôs, redes sociais e política no Brasil [recurso eletrônico]: estudo sobre interferências ilegítimas no debate público na web, riscos à democracia e processo eleitoral de 2018. Rio de Janeiro: FGV, DAPP, 2017.

RUEDIGER, Marco Aurélio (Coord.). Robôs, redes sociais e política no Brasil [recurso eletrônico]: casos de interferências ilegítimas no debate público por automação de perfis. Rio de Janeiro: FGV, DAPP, 2018.

SANTOS, Maíke Wile dos. O big data somos nós: a humanidade de nossos dados. *JOTA*, 16 mar. 2017. Disponível em: <<https://jota.info/colunas/agenda-da-privacidade-e-da-protecao-de-dados/o-big-data-somos-nos-a-humanidade-de-nossos-dados-16032017>>. Acesso em: 27 mar. 2017.

SILVERMAN, Craig. Here Are 50 Of The Biggest Fake News Hits On Facebook From 2016. *BuzzFeed News*, dez. 2016. Disponível em: <<https://www.buzzfeednews.com/article/craigsilverman/top-fake-news-of-2016#.nl712lkw2>>. Acesso em: 29 out. 2018.

SIMÕES, Mariana. Grupos pró-Bolsonaro no WhatsApp orquestram notícias falsas e ataques pessoais na internet, diz pesquisa. *El País*, out. 2018. Disponível em: <[https://brasil.elpais.com/brasil/2018/10/23/politica/1540304695\\_112075.html?id\\_externo\\_rsoc=FB\\_BR\\_CM&fbclid=IwAR05Mw9zXzmjDbYv5OkjAmrhVipWBURMCPyiOORlaxSsy\\_qNxEjzrpHKxfQ](https://brasil.elpais.com/brasil/2018/10/23/politica/1540304695_112075.html?id_externo_rsoc=FB_BR_CM&fbclid=IwAR05Mw9zXzmjDbYv5OkjAmrhVipWBURMCPyiOORlaxSsy_qNxEjzrpHKxfQ)>. Acesso em: 29 out. 2018.

SUNSTEIN, Cass. *Republic.com 2.0*. Princeton: Princeton University Press, 2009.

\_\_\_\_\_. *Republic.com*. Princeton and Oxford: Princeton University Press, 2001.

VELASCO, Clara; DOMINGOS, Roney. O que é um robô na web e como ele pode influenciar o debate nas redes? Especialistas explicam. *GI*, dez. 2017. Disponível em: <<https://gi.globo.com/economia/tecnologia/noticia/o-que-e-um-robo-na-web-e-como-ele-pode-influenciar-o-debate-nas-redes-especialistas-explicam.ghtml>> Acesso em: 29 out. 2017.

VIANA, Natalia; ZANATTA, Carolina. Deep fakes são ameaça no horizonte, mas ainda não são arma para eleições, diz especialista. *A Pública*, out. 2018. Disponível em: <<https://apublica.org/2018/10/deep-fakes-sao-ameaca-no-horizonte-mas-ainda-nao-sao-arma-para-eleicoes-diz-especialista>>. Acesso em: 25 out. 2018.

WIGGERSHAUS, Rolf, et al. *The Frankfurt School: Its History, Theories, and Political Significance*. The MIT Press, 1995.

WU, Tim. *The Master Switch: The Rise and Fall of Information Empires*. New York: Vintage, 2011.