

Estado da publicação: Não informado pelo autor submissor

Fake News: O Que É, Como Se Faz E Por Que Funciona?

Renê Forster, Rodrigo Monteiro de Carvalho, Alberto Filgueiras, Emanuelle Avila

<https://doi.org/10.1590/SciELOPreprints.3294>

Submetido em: 2021-11-27

Postado em: 2021-12-07 (versão 1)

(AAAA-MM-DD)

Fake News: O Que É, Como Se Faz E Por Que Funciona?

Fake News: What Is It, How Is It Made and Why It Works?

Renê Forster

Departamento de Estudos da Linguagem – Universidade do Estado do Rio de Janeiro
<https://orcid.org/0000-0002-3681-1505>

Rodrigo Monteiro de Carvalho

Doutorando em Estudos Contemporâneos pelo Centro de Estudos Interdisciplinares do Século XX (CEIS20), da Universidade de Coimbra
<https://orcid.org/0000-0002-0263-7692>

Alberto Filgueiras

Departamento de Cognição e Desenvolvimento – Universidade do Estado do Rio de Janeiro
<https://orcid.org/0000-0002-6668-0606>

Emanuelle Avila

Graduanda em Letras - Universidade do Estado do Rio de Janeiro
<https://orcid.org/0000-0002-9370-6541>

RESUMO

Este trabalho apresenta uma revisão sobre a desordem informacional, incluindo as chamadas “fake news”. Conteúdos falsos ou maliciosos, por seu potencial de propagação nos meios digitais, impactam a vida política e social, sendo associados desde a rejeição de recomendações de saúde até a distúrbios violentos. Nessa revisão, em primeiro lugar, caracterizamos a desordem informacional, elencando distinções conceituais entre suas diferentes manifestações. Em segundo lugar, tratamos dos processos produção e divulgação desses conteúdos, abordando os agentes e canais envolvidos. Na sequência, debatemos o processo de recepção de conteúdos falsos. São apresentados trabalhos que, a partir de mecanismos de ordem cognitiva e psicológica, buscam responder por que se acredita e se socializa esses materiais.

ABSTRACT

This work presents a review about information disorder, amidst which we include the so-called “fake news” phenomenon. Due to its increased potential for propagation in digital age, false or malicious content has impacted political and social life and its consequences range from rejection of health recommendations to violent disorders. In this review, we first define information disorder, considering some theoretical distinctions between its different manifestations. We then address the processes of production and dissemination of false and malicious content, focusing the producers and also the channels through which such materials are spread. Next, we discuss reception of false content, reviewing works on the psychological and cognitive bases of gullibility.

RESUMEN

Este trabajo presenta una revisión del desorden informativo, incluyendo las llamadas “fake news”. Los contenidos falsos o maliciosos, por su potencial de propagación en los medios digitales, impactan la vida política y social, asociándose desde el rechazo de recomendaciones de salud a disturbios violentos. En esta revisión, en primer lugar, caracterizaremos el desorden informativo, enumerando las distinciones conceptuales entre sus diferentes manifestaciones. En segundo lugar, nos ocuparemos de los procesos de producción y difusión de esos contenidos, planteando los agentes y canales implicados. A continuación, discutiremos el proceso de recibo de los contenidos falsos. Se presentan trabajos que, a partir de mecanismos de orden cognitivo y psicológico, buscan responder por qué se cree y socializan esos materiales.

RÉSUMÉ

Ce travail présente une révision sur le désordre d'information, considérant ce qu'on appelle "fake news". Des contenus faux ou malins, d'après son potentiel de diffusion dans les milieux numériques, ont un impact sur la vie politique et sociale, étant associés au refus de recommandations de santé jusqu'au troubles violents. Lors de cette révision, dans un premier moment, on caractérise le désordre d'information, en énumérant des disparités de concepts entre ses différentes manifestations. Dans un deuxième moment, on s'occupe des processus de production et diffusion de ces contenus, traitant les agents et les voies impliquées. Après, on discute le processus d'accueil des faux contenus. Seront présentés des travaux qui, d'après les mécanismes d'ordre cognitive et psychologique, essaient de répondre pourquoi on croit et on diffuse ces matériaux.

PALAVRAS-CHAVE: desordem informacional; desinformação; fake news; notícias falsas; mídias sociais, viés confirmatório

KEYWORDS: information disorder; disinformation; fake news; social networks; confirmation bias

1. INTRODUÇÃO

A disputa pelo poder entre dois rivais é assolada pela divulgação de notícias falsas que acabam por prejudicar a reputação de um dos oponentes e facilitar a ascensão do outro. Embora nos seja uma história familiar, esse episódio aconteceu na década de 30 a.C., quando Otaviano, posteriormente proclamado Imperador Augusto, lançou mão de uma campanha de difamação contra seu maior rival pelo poder em Roma, Marco Antônio. Poesias lidas em público, pequenos textos que passavam de mão em mão e inscrições cunhadas em moedas serviram de instrumento para retratar Antônio como um bêbado irresponsável e um adúltero (KAMINSKA, 2018). Esses artifícios, que se provaram eficientes à época, foram se reciclando ao longo dos séculos até chegarem na era digital (POSETI e MATTHEWS, 2018; MAGALHÃES, 2018; SOLL, 2016). O distópico mercado das "fake news", tratado inicialmente com certo assombro e surpresa, hoje tem sido entendido como produto da evolução de ferramentas tecnológicas associado a circunstâncias sociais, políticas e até mesmo a mecanismos de ordem cognitiva.

No domínio tecnológico, a internet, as redes sociais e os aplicativos de mensagens revolucionaram as formas pelas quais materiais falsos ou maliciosos podem se disseminar. Conteúdos podem ser replicados, sem intermediários, com rapidez, em grande volume e até mesmo de forma anônima. As baixas barreiras para a produção e compartilhamento de informação e a capilaridade dos recursos de tecnologia da informação fizeram do ambiente digital o centro da arena política na contemporaneidade (Cf. CASTELLS, 2004; CHADWICK e HOWARD, 2010; JENKINS, 2016; FARKAS e SCHOU, 2019).

Mas não foi apenas a tecnologia que permitiu a escalada das "fake news". A falta de políticas e normas claras direcionados ao mercado de tecnologia da informação, a concentração das ferramentas populares nas mãos de poucos e a escassez de políticas educacionais específicas para a era digital estão dentre os fatores que fizeram nascer toda uma indústria de informações falsas ou maliciosas. Hoje, a produção de conteúdos caluniosos, difamatórios e inverídicos é vendida por empresas especializadas e sua distribuição pode ser contratada nas plataformas sociais. Esse ecossistema gera lucros a produtores e intermediários, permitindo que diversos atores (de políticos a empresas) construam narrativas e defendam agendas com base em informações falsas ou mal-intencionadas (BRADSHAW et al., 2020; BRADSHAW e HOWARD, 2019; GU et al., 2017; DE CRISTOFARO et al., 2014).

Na outra ponta, os consumidores de informação, com a atenção desorientada pelo volume de conteúdos, são levados a processar o que recebem de forma acrítica. Isso

acontece, em parte, em razão de tendências cognitivas que nos tornam menos vigilantes a incongruências e nos levam a confiar em informação com base em pistas pouco consistentes, como a repetição exaustiva, a presença da fonte em nossa rede de contatos ou seu alinhamento ao nosso sistema de crenças (PENNYCOOK et al., 2015; JAGATIC et al., 2007; BUCHANAN e BENSON, 2019; FREINDA et al., 2013; WANG et al. 2020).

Os impactos desse circuito de conteúdos falsos e maliciosos já são amplamente documentados: danos à saúde pública, violência étnico-política e perturbação de processos eleitorais. A propagação descontrolada de conteúdos inverídicos ou mal-intencionados é considerada hoje uma ameaça às democracias.

Neste artigo, apresentamos uma revisão sobre o debate da desordem informacional. O principal objetivo deste trabalho é oferecer um material introdutório ao tema, numa perspectiva multidisciplinar, que considera desde os mecanismos socioeconômicos aos fatores psicológicos envolvidos no fenômeno e que leva em conta os estudos acadêmicos e a cobertura jornalística sobre o tema. Embora materiais introdutórios já possam ser encontrados (como os publicados, por exemplo, pelo Global Desinformation Index¹ e pelo Internet Lab²), a literatura em língua portuguesa ainda parece carecer de um material dessa natureza em formato de artigo.

Como ponto de partida, buscamos discutir alguns aspectos terminológicos. Embora termos como “fake news” sejam utilizados frequentemente tanto na imprensa como na literatura especializada, parece ainda não haver um consenso sobre as categorias de informação às quais esse termo se referiria. Nessa discussão, tentamos apresentar algumas alternativas mais consistentes ao termo, além de outras definições análogas. Na terceira seção do artigo, revemos trabalhos que investigam como as informações falsas ou maliciosas são produzidas e disseminadas. Abordamos tanto os atores quanto os meios envolvidos na confecção e circulação desses conteúdos, procurando demonstrar em particular como a industrialização desse processo permitiu que a desordem informacional se tornasse um componente sistêmico do funcionamento das sociedades contemporâneas. Na quarta seção, abordamos o problema pelo prisma de sua recepção, com o objetivo de entender como somos levados a acreditar em conteúdos falsos. Nessa direção, apresentamos trabalhos que discutem diversos mecanismos psicológicos envolvidos na assimilação de informação. Na seção final do artigo, são considerados também prognósticos para o cenário da desordem informacional.

2. O QUE SÃO AS FAKE NEWS? FENÔMENOS DA DESORDEM INFORMACIONAL

2.1 A desordem informacional

Há pouca consistência terminológica na literatura que trata de notícias falsas e fenômenos relacionados. Algumas tentativas de se criar um arcabouço conceitual para sanar esse problema foram empreendidas nesses últimos anos³ e o relatório “*Information Disorder: Toward an interdisciplinary framework for research and policymaking*” (WARDLE e DERAKHSHAN, 2017b) encomendado pelo Conselho Europeu em 2017, se tornou um trabalho referência desse esforço. Nele é apresentado o conceito de “desordem informacional”, que abrange três fenômenos: *misinformation*, a veiculação não intencional de informações equivocadas; *malinformation*, o uso de informações verdadeiras, mas manipuladas de modo a enganar; e *disinformation*, que serve como um grande guarda-chuva para técnicas intencionais de disseminação de informações necessariamente falsas e com o objetivo específico de criar confusão ou causar danos premeditados (WARDLE e DERAKHSHAN, 2017b).

2.1.1 Misinformation

¹ <https://disinformationindex.org/>

² <https://www.internetlab.org.br/pt/>

³ Ver GELFERT, 2018; JACK, 2017; TANDOC JR, LIM, e LING, 2018.

A *misinformation* se caracteriza pela não intencionalidade de desinformar. Vale tanto para quem produz uma notícia inverídica por descuido, como para quem, sem dolo, as dissemina. Em 2014, diversos veículos de imprensa internacionais noticiaram que o governo norte-coreano fez os seus cidadãos acreditarem que a seleção nacional teria conquistado três vitórias consecutivas na Copa do Mundo disputada no Brasil, torneio para qual a Coreia do Norte nem havia se classificado (PECK, 2014). Tratava-se de uma paródia que acabou por ser veiculada como verdade. Dessa forma, é possível considerar que a informação foi publicada por ausência de adequada apuração jornalística. Isto, *a priori*, não caracteriza má fé, portanto um caso de *misinformation*. A recorrência de episódios como esse no meio jornalístico é agravada pela costumeira ausência de princípios claros de apuração, o que acaba por gerar uma influência demasiada dos tópicos discutidos em redes sociais nas decisões editoriais da imprensa (ADORNATO, 2016). Pela falta de intenção de causar dano ser o elemento definitivo para a caracterização da *misinformation*, e julgar intenção não ser tarefa trivial, resta uma dificuldade patente de se classificarem esses casos.

2.1.2 Malinformation

A *malinformation* se assenta no campo da reconfiguração da verdade, ou seja, faz uso de elementos verdadeiros para construir uma narrativa com o intuito de manipular opiniões e promover agendas específicas. Estão nesse âmbito o sensacionalismo, a descontextualização; os discursos de ódio (aqueles que se valem de informações reais, mas reinterpretadas de modo a incitar a revolta contra pessoas ou grupos); o *clickbait* (manchetes, legendas ou imagens que não condizem com o conteúdo); e alguns tipos de vazamentos de dados sigilosos, especificamente aqueles com o intuito de causar prejuízos ou auferir ganhos, em geral políticos (como os vazamentos de delações contra candidatos às vésperas de eleições).

De forma geral, o que vem alavancando a disseminação da *malinformation* é o crescente número de veículos politicamente engajados, como os sites ou blogs “hiperpartidários” (MARWICK e LEWIS, 2017, p. 21; SILVERMAN et al., 2017), direcionados a fomentar o discurso de ódio, o acirramento das convicções ideológicas e o aprofundamento da polarização. Esse tipo de veículo pode publicar tanto informação simplesmente falsa, como também dados verdadeiros, mas apresentados de modo a promover o medo, a raiva e sedimentar a radicalização de grupos opositores. Essa prática é classificada por ORTELLADO e RIBEIRO (2018) como “informação de combate”.

2.1.3 Desinformação

A desinformação, por sua vez, se caracteriza pelo uso de informações falsas com a intenção de causar dano. Trata-se então de conteúdos falsos e maliciosos. Há inúmeros exemplos recentes, desde aqueles em que se busca algum grau de verossimilhança com uma informação real (como a que alega que a aposentadoria precoce de Jair Bolsonaro do Exército se deu por invalidez psicológica (DOMINGOS, 2018)) até conteúdos de teor absurdo (como o alerta de que antenas de transmissão 5G estariam ligadas à disseminação da Covid-19 (SATARIANO e ALBA, 2020)). O recurso a informações falsas e maliciosas com fins políticos tem longa trajetória e marcou diversos processos eleitorais no Brasil, como no caso dos “marmiteiros” ainda em 1945, passando pelo emblemático caso das denúncias contra Lula no debate eleitoral de 1989 (OBSERVATÓRIO DA IMPRENSA, 2012; SAKAMOTO, 2016; MAGALHÃES, 2018).

O termo “desinformação” pode se referir a diversos formatos, desde conteúdo jornalístico fabricado, passando por memes, vídeos, imagens, enquetes digitais e petições online (para uma apresentação de subtipos de desinformação e *misinformation*, veja-se WARDLE (2019) e WARDLE e DERAKSHAN, (2018)).

2.1.4 Fake news

É no âmbito das manifestações da desinformação que mais tipicamente são situadas as “fake news”. A expressão “fake news” se popularizou em meio à campanha presidencial estadunidense de 2016 e desde então é difícil encontrar alguém que não a atribua algum significado. Mas embora ubíquo, e sobretudo por isso, o termo passou a mais confundir do que explicar. Primeiro porque o termo muitas vezes é usado em referência a informações maliciosas que não são necessariamente transmitidas em formato de notícia jornalística *stricto sensu*. Assim, o termo parece acabar abrangendo também diversos formatos de conteúdo, como, por exemplo, memes difamatórios ou maliciosos. Ademais, “fake news” passou a ser expressão usada para classificar toda informação, verdadeira ou não, que desagrade ou vá de encontro às convicções já formadas por alguém. O grande popularizador dessa última tendência foi o Presidente estadunidense Donald Trump, que frequentemente rotula seus críticos como sendo propagadores de “fake news” (SCHWARTZ, 2018). Embora um dos precursores, Trump não foi o único a adotar a tática de usar o conceito a seu favor (LEES, 2018; VISCARDI, 2020).

Existe, além disso, certa inconsistência quanto ao aspecto da intencionalidade: enquanto autores restringem o termo a conteúdo noticioso produzido maliciosamente, outros usam o termo em referência a qualquer conteúdo noticioso falso seja ele intencionalmente produzido ou não (Cf. seção 2.4). A imprecisão do termo “fake news” levou o Reino Unido a banir a expressão de seus documentos oficiais (PALMER, 2018)⁴.

É importante notar que, embora as notícias falsas propriamente ditas mereçam preocupação, é a desinformação transmitida por suporte audiovisual (memes contendo informações falsas ou imagens, áudios e vídeos adulterados) o tipo de conteúdo mais pervasivos (MOSCHELLA e WATTS, 2017). Desse modo, é frequente que uma campanha desinformativa seja reforçada pela combinação de mais de um artifício, de modo que se amplifique o alcance e a penetração da mensagem.

2.1.5 Outros fenômenos da desordem informacional

Cabe ainda acrescentar que, embora as três categorias de desordem informacional deem conta da maior parte dos fenômenos existentes, há alguns pontos cegos que merecem destaque. É o caso da propaganda e das teorias conspiratórias. A propaganda, sobretudo aquela empreendida por governos, costumeiramente emprega elementos da *malinformation* e da desinformação (Cf. JACK, 2017). As teorias conspiratórias que, por definição, não estão sustentadas em fatos comprováveis, também parecem não ser contempladas por essa tipologia. Já outras manifestações, como alguns tipos de *clickbait* e de descontextualização (classificadas tipicamente como *malinformation* por apresentarem elementos verdadeiros em seu conteúdo) podem ser alternativamente entendidas como desinformação. Isso se dá porque nem sempre é claro o limiar entre o que são informações verdadeiras rearrumadas de modo a servir propósitos específicos e o que são informações indubitavelmente falsas construídas com uso de alguns poucos elementos reais.

Um outro termo recorrente na literatura é “propaganda computacional” (*computational propaganda*), em referência às táticas de manipulação da opinião pública em ambiente digital, a serviço de grupos políticos ou governos, que recorrem ao uso de automação e algoritmos, em especial, nas redes sociais (WOOLEY e HOWARD, 2017; BRADSHAW e HOWARD, 2018). O escândalo da Cambridge Analytica (Cf. seção 3.1), por exemplo, seria uma das manifestações desse tipo de estratégia. O uso do termo em questão deixa a entender que o rótulo se refere não a tipos de conteúdos, mas a práticas de distribuição de informação, que tipicamente envolvem conteúdos falsos, mas que parecem poder envolver conteúdos verdadeiros também. Embora a prática de propaganda computacional suponha uma intencionalidade, os conteúdos envolvidos nessa modalidade de propaganda podem ou não ser produzidos de maneira intencional.

⁴ Para uma discussão acerca do uso dos termos “fake news” e “pós-verdade”, veja-se HABGOOD-COOTE (2018).

2.2 Grande mídia

É necessário considerar também que, muito embora a desordem informacional seja tipicamente situada no âmbito da produção de conteúdos fora do espectro dos agentes tradicionais e/ou corporativos de informação, ou seja, a chamada grande imprensa ou grande mídia, há que se ter também a perspectiva de que esses agentes, muitas vezes atuando em prol de agendas e interesses específicos, podem lançar mão de práticas de *malinformation*, distorcendo, manipulando ou enviesando conteúdos verdadeiros (para uma discussão sobre o funcionamento da mídia tradicional e seu alinhamento a agendas específicas, veja-se, por exemplo, HERMAN e CHOMSKY, 1988 e CHOMSKY, 2015). Nesse sentido, há um grande debate, por exemplo, acerca do papel da imprensa na cobertura do caso das armas de destruição em massa que teriam levado os Estados Unidos e o Reino Unido a declararem guerra ao Iraque, em 2003 (MOELLER, 2004; SHWARZ, 2015).

2.3 A cadeia da informação

A cadeia da informação pode ser subdividida em três fases: a concepção, a produção e a distribuição (WARDLE e DERAKHSHAN, 2017b, p. 22). A fase de concepção é o momento em que a informação é idealizada e são definidos os objetivos que devem ser atingidos com sua propagação. A produção se dá quando a ideia é transformada em uma peça informacional pronta para ser posta pública. Já a distribuição é a etapa final de disseminação do conteúdo. Um agente único pode controlar toda a cadeia, da mesma forma que agentes diferentes podem ser os responsáveis por cada uma das etapas. Nesse sentido, um idealizador/financiador de uma campanha de desinformação pode delegar a tarefa de produção e distribuição para uma miríade de agentes diferentes.

Dadas essas diferentes etapas, torna-se possível que mais de uma prática seja combinada durante o processo de construção da cadeia da informação. Imaginemos, como exemplo, que uma informação falsa criada intencionalmente para causar danos a um dado candidato em uma eleição, portanto, um caso de desinformação. Mas essa notícia chega a um eleitor que a toma como verdade e a espalha entre seus amigos, ou seja, um ato de *misinformation*. Em sentido inverso, é possível conceber que uma notícia falsa construída sob preceitos equivocados, mas sem a intenção de causar algum dano específico, pode alcançar um intermediário que, mesmo sabendo se tratar de uma inverdade, decide, em prol de algum interesse, disseminá-la, portanto, um ato de desinformação.

2.4 Terminologia adotada

No presente trabalho, procuramos nos orientar pela terminologia sugerida por WARDLE e DERAKHSHAN (2017b) e apresentada ao longo desta seção. Vale, no entanto, considerar algumas situações em particular. Nos casos em que não se determina a intencionalidade de um conteúdo falso, usamos os termos conteúdos, materiais ou informações falsas. Esses termos podem englobar então tanto o que se chama de *misinformation* quanto o que se chama de desinformação. Note-se que, em alguns dos estudos que revisamos neste texto, essa distinção deixa inclusive de ser pertinente. A título de exemplo: em alguns dos trabalhos apresentados na seção 4, os conteúdos falsos são estímulos experimentais produzidos com fins de pesquisa. A intencionalidade na produção destes materiais é, portanto, óbvia. Contudo, uma vez que essa intenção não advém do propósito de causar dano ou auferir ganhos políticos ou materiais, devemos considerá-la de maneira diferente da intencionalidade relacionada ao rótulo de “desinformação”.

No caso do termo notícias falsas, buscamos adotar o enquadramento sugerido por Lazer e colegas (2018): “we define “fake news” to be fabricated information that mimics news media content in form but not in organizational process or intent”, segundo o qual trata-se de conteúdo noticioso falso, produzido com ou sem intencionalidade. Uma vez que o presente trabalho se trata de uma revisão, consideramos esta definição em particular, uma vez que muitos dos estudos analisados usam o termo passando ao largo da questão da

intencionalidade. Vale notar, contudo, que uma definição mais estrita, tal qual a aceção de ALLCOTT e GENTZKOW (2017, p. 213), também é assumida por alguns pesquisadores: “*artigos noticiosos que são intencional e comprovadamente falsos, e podem enganar os leitores*”⁵. Nesse sentido, o termo faria referência a conteúdo noticioso falso produzido intencionalmente.

O termo “fake news”, em inglês, embora apresentado no título e na introdução por sua relevância política, é, no restante do texto, utilizado mais particularmente quando, metalinguisticamente, nos referimos à própria terminologia.

Vale notar, por fim, que embora tenhamos tentado preservar certa consistência terminológica, em alguns casos nosso relato poderá refletir a vagueza, imprecisão ou inconsistência dos estudos analisados, derivados muitas vezes de algumas dificuldades de classificação já apontadas nesta seção.

3. PRODUÇÃO E DIVULGAÇÃO: COMO SURGE A DESINFORMAÇÃO?

Nesta seção, apresentamos uma análise da produção e circulação de conteúdos falsos, com ênfase na desinformação. Nesse sentido, buscamos, por um lado, (a) elencar alguns dos agentes envolvidos nesse processo e, por outro, (b) caracterizar as plataformas nas quais esses materiais se originam e circulam.

3.1 Atores da produção e circulação de desinformação

O processo de produção de desinformação atualmente é uma dinâmica que mescla atividades institucionalizadas e clandestinas. A face institucionalizada dessa operação permite, por um lado, o financiamento de produtores de desinformação em ambientes nos quais, por exemplo, um CNPJ seja requerido para transações financeiras. A clandestinidade dessas atividades, por sua vez, permite que os limites éticos ou morais sejam escamoteados, impõe dificuldades à responsabilização e permite ainda a defesa de agendas extremistas sem o desgaste de seus promotores. Essa produção pode ser motivada por impulsos narcisistas, como a busca por status ou atenção (LEWIS e MARWICK, 2017; MARWICK e LEWIS, 2017; PAPACHARISSI, 2010). Contudo, a motivação que parece dar sistematicidade a essas atividades (de maneira não necessariamente dissociada das demais motivações) está na geração de resultados financeiros ou políticos.

A possibilidade de gerar renda com a desinformação faz com que até mesmo atores totalmente alheios aos impactos políticos concretos desse material se envolvam em sua produção. Sites ativos na produção de notícias durante a campanha eleitoral norte-americana de 2016, por exemplo, eram gerenciados da Macêdonia. Notícias apelativas direcionadas, em particular, ao eleitorado de Donald Trump propiciavam renda de cliques patrocinados (Cf. seção 3.2) num sistema mercenário de produção de desinformação (SILVERMAN, 2016; SUBRAMANIAN, 2017).

Mas a desinformação pode ser patrocinada também por aqueles diretamente interessados em seus resultados. Governos, partidos e políticos, buscando a construção ou a sustentação de narrativas, financiam e promovem sistematicamente a desinformação. Há evidência de que a manipulação de redes sociais por governos, por exemplo, ocorra em 81 países, segundo dados de 2020 (BRADSHAW et al., 2021). No caso de uso por políticos, o exemplo mais emblemático é a campanha presidencial norte-americana de 2016. Numa prática conhecida com *microtargeting*, desinformação teria sido distribuída com base no perfilamento de eleitores. A organização responsável por esta estratégia, Cambridge Analytica, teria atuado em mais de 200 eleições (POSETTI e MATTHEWS, 2018). No Brasil, há evidência de que parlamentares, partidos e organizações políticas estejam diretamente envolvidos no financiamento de práticas de desinformação (LOCATELLI, 2018; TOLEDO, 2018; ESTADÃO, 2020a; NEMER, 2020). Gabinetes de deputados estaduais, federais e até

⁵ “We define “fake news” to be news articles that are intentionally and verifiably false, and could mislead readers” (ALLCOTT e GENTZKOW, 2017, p. 213).

da presidência da república já foram associados ao compartilhamento de discurso de ódio e desinformação em redes sociais, segundo investigação interna do Facebook (GLEICHER, 2020).

Para que o financiamento de conteúdo falso possa ser feito com dinheiro público ou corporativo, seus produtores precisam se adequar a requerimentos burocráticos. Se desenvolve então o empresariamento de atividades de desinformação. Um exemplo paradigmático nesse sentido é a Novo Brasil Empreendimentos Digitais, responsável pelo conglomerado de desinformação Folha Política. Formalmente constituída, a empresa conseguiu captar recursos tanto de partidos políticos quanto de cotas parlamentares (LOCATELLI, 2018; TOLEDO, 2018). Outro exemplo é o Jornal da Cidade Online⁶, nome fantasia da A. J. Pinheiro Tolentino Filho EIRELI, também associado a publicação de desinformação. A empresa foi capaz de captar, segundo dados do Tribunal de Contas da União, recursos de publicidade oriundos de empresas estatais via plataformas de anúncios pagos (ESTADÃO, 2020b; VALOR ECONÔMICO, 2020). Apesar do patrocínio ter se dado por uma via indireta, O Banco do Brasil, uma das estatais envolvidas, defendeu a exposição de publicidade no Jornal da Cidade Online (BANCO DO BRASIL, 2020), o que sugere a hipótese, que discutiremos a seguir, de que as plataformas de anúncios pagos sejam usadas com a finalidade de dissimular vínculos entre patrocinadores e produtores de desinformação.

3.2 Blogs e sites: cliques pagos

No universo da desordem informacional, blogs e sites funcionam, muitas vezes, como fontes primárias de conteúdos falsos ou maliciosos, em particular, de material noticioso. Embora esses espaços possam parecer amadores, há a profissionalização da operação desses canais, o que é traduzido, muitas vezes, em retornos financeiros para seus administradores. Considerando o nosso propósito de destacar a sistematicidade dos mecanismos dessa produção, nesta seção, abordamos os blogs e sites a partir de suas estratégias de financiamento. Em particular, apresentamos três dessas estratégias.

Uma das fontes de financiamento são as plataformas de links patrocinados como o Google Ads e o Taboola. Essas plataformas possibilitam a inserção de espaços de publicidade em endereços da web. Esses espaços remuneram quem os hospeda pelos cliques nos banners de anúncios. Uma vez que as plataformas de anúncios oferecem poucas barreiras para a inserção desses espaços publicitários, esse mecanismo, então, permite que produtores de conteúdos maliciosos possam gerar renda. Como exemplo da lucratividade desse sistema, há a estimativa de que, a cada quatro dólares arrecadados por sites com informações falsas sobre a Covid-19, três fossem advindos do Google Ads (TOGOH, 2020; ver também TAYLOR et al., 2020). Os anunciantes, por sua vez, desfrutam da possibilidade de, a partir das informações que essas plataformas coletam sobre os perfis de navegantes, direcionar seus produtos de forma customizada para seu público-alvo. A chave do funcionamento desse sistema para o patrocínio da desordem informacional está na possibilidade de que os anunciantes, uma vez que não contratam diretamente os criadores de conteúdo, possam alegar não ter vínculos com esses produtores (eg. TOGOH, 2020; BANCO DO BRASIL, 2020; BRADSHAW et al., 2021). A camuflagem dos vínculos operada pelas plataformas de anúncios é, então, o elemento essencial para que a desinformação seja patrocinada nesse sistema. Os anunciantes, contudo, não deixam de se beneficiar da atratividade desses conteúdos. Há evidência de que essa estratégia de captação de recursos teria financiado, por exemplo, importantes sites de desinformação como Diário do Brasil⁷ e Jornal do País⁸ (Cf. seção 3.1), que atuaram incisivamente em processos políticos no país (GHEDIN et al., 2019).

Outra estratégia para obter lucros a partir de conteúdo falso ou malicioso envolve a repercussão de matérias de blogs e sites em outras plataformas monetizadas. Nesse caso, a matéria publicada em um endereço da web serve principalmente à finalidade de direcionar o

⁶ <https://www.jornaldacidadeonline.com.br/>

⁷ <https://www.diariodobrasil.org/>

⁸ <https://www.jornaldopais.com.br/>

leitor para um meio no qual o produtor de conteúdo poderá se beneficiar financeiramente da audiência. O YouTube, por ser uma das ferramentas que paga aos produtores por visualização de vídeos, é um dos instrumentos utilizados nesse sentido. Um exemplar dessa estratégia é o canal Folha Política⁹, possivelmente um dos mais representativos sites de notícias falsas (ORTELLADO e RIBEIRO, 2018; FARAH, 2018). Em maio de 2020, este canal contava com quase 2 milhões de inscritos, tendo com isso um potencial de geração mínima de 10.400 dólares por mês na plataforma audiovisual segundo o site de análise de tráfego Social Blade¹⁰. Atualmente, a coluna central do site é integralmente dedicada a matérias que direcionam para vídeos no YouTube, o que sugere a importância dessa estratégia de monetização. A relevância do mercado de cliques pagos como instrumento de financiamento da desinformação tem levado, inclusive, ao surgimento de movimentos de mobilização que denunciam essa prática (VEIGA, 2020).

Uma terceira estratégia de monetização é o patrocínio direto aos sites. Nesse caso, atores interessados nos conteúdos e suas repercussões pagam diretamente aos produtores, em geral sob a camuflagem de serviços de publicidade. Esse sistema evidencia, em especial, o empresariamento da produção de notícias falsas. Um exemplar dessa estratégia é o Notícias Naturais¹¹. Em 2019, o Notícias Naturais foi elencado como o principal produtor de conteúdo antivacina no Brasil (AVAAZ e SBIM, 2019). O site, articulado a páginas no Facebook, contava com um anunciante fixo e único, a loja online de produtos naturais Tudo Saudável¹². A partir de conteúdos sensacionalistas sobre saúde, como “Melhores Alimentos para Queimar a Gordura da Barriga”¹³ ou “Vacinação contra HPV: Conheça o perigo que ela representa e tome uma decisão informada”¹⁴, o site atraía leitores potencialmente interessados em tratamentos alternativos e produtos naturais (AVAAZ e SBIM, 2019). Outro exemplo de financiamento direto é o patrocínio de sites de desinformação por agentes políticos, discutido na seção anterior (Cf. seção 3.1).

Tanto pelo mercado de cliques quanto pelo financiamento direto, é possível observar que informações falsas ou maliciosas não beneficiam apenas seus produtores. Políticos e empresas se favorecem desse mercado. Há inclusive indícios de que o Google, por exemplo, ofereceria treinamento a produtores de notícias falsas com o objetivo de alavancar sua plataforma de anúncios no Brasil (GHEDIN et al., 2019). Embora parte da produção de desinformação possa ser clandestina, é necessário que seus produtores, a fim de se financiar e crescer neste mercado, recorram à institucionalização, como se vê pelo empresariamento dessa produção e por seu financiamento com recursos públicos (Cf. seção 3.1).

3.3 Redes sociais

A circulação de informações falsas ou maliciosas depende, em grande medida, das redes sociais, de modo que produtores desse tipo de material costumam ter presença coordenada em diversas plataformas (ORTELLADO e RIBEIRO, 2018). Essas redes reduziram as barreiras tanto para a publicação quanto para a divulgação de informações e, além disso, se tornaram o principal veículo de consumo de notícias (NEWMAN et al., 2019). A conjunção desses dois fatores torna as redes sociais um terreno fértil para a propagação de informações, sejam elas verdadeiras ou não. Ademais, a engenharia dessas plataformas e os comportamentos que elas incentivam fazem com que essas plataformas tenham se tornado o epicentro da desordem informacional.

⁹ <https://www.youtube.com/channel/UCYiM773ssvNMaBHvaWWeloQ>

¹⁰ <https://socialblade.com/youtube/channel/UCYiM773ssvNMaBHvaWWeloQ>

¹¹ www.noticiasnaturais.com (até abril de 2020)

¹² <https://www.tudosaudavel.com/>

¹³ <https://web.archive.org/web/20191024063403/https://www.noticiasnaturais.com/2019/10/7-melhores-alimentos-para-queimar-a-gordura-da-barriga/#>

¹⁴ <https://web.archive.org/web/20191024063005/https://www.noticiasnaturais.com/2014/03/dossie-vacinacao-contrahpv-conheca-o-perigo-que-ela-representa-e-tome-uma-decisao-informada/>

Nessa seção, consideramos alguns aspectos da engenharia das redes sociais e sua relação com as dinâmicas de produção e circulação de informações falsas ou maliciosas. Para essa análise, consideramos, em parte, o trabalho de Kietzmann et al. (2011), que elenca alguns “elementos funcionais” das mídias sociais.¹⁵ A partir deles, procuramos demonstrar que a desordem informacional é favorecida pela arquitetura das redes sociais. Embora não se possa afirmar que tal favorecimento seja intencional, ao mesmo tempo, será possível perceber, em nossa análise, que as práticas de desinformação são detectáveis em todas as principais redes sociais. Além disso, não se pode negar, diante de todas as evidências publicizadas até hoje, que essas práticas sejam de conhecimento das companhias por trás dessas plataformas.

3.3.1 Anonimato e Manipulação de identidades

Uma das características da comunicação em âmbito digital é a de que as mensagens ou publicações nesse universo e, em particular, nas mídias sociais não estão necessariamente vinculadas a uma identidade rastreável ou responsabilizável. Embora os termos de serviço de algumas plataformas, como o Facebook, proíbam o uso de identidades falsas (FACEBOOK, 2020), na prática, seu uso não é incomum e é, algumas vezes, assimilado nos próprios termos da rede, como no caso do Twitter (TWITTER, 2020).

Com base nessa possibilidade, surge todo um mercado de perfis sociais fraudulentos, que podem até ser comprados no atacado (THOMAS et al., 2013). O anonimato ou a criação de proxys é um recurso que acaba permitindo a manipulação dos algoritmos das redes sociais. Em redes como Facebook, Instagram e Twitter, o conteúdo é apresentado aos usuários na forma de uma “linha do tempo” (do inglês, *timeline*). Esses conteúdos são ranqueados por um algoritmo, de modo que um usuário veja prioritariamente publicações bem classificadas. As interações (eg., curtidas, comentários, compartilhamentos) despertadas por um determinado conteúdo são um dos fatores que impulsionarão seu ranqueamento. As baixas barreiras para a criação de perfis e as possibilidades de dissimulação de identidade permitirão o desenvolvimento de estratégias de manipulação das interações e, conseqüentemente, de manipulação deste ranqueamento.

Parte das estratégias nesse sentido recorre à utilização de robôs, ou em outros termos, bots sociais. Trata-se de contas parcialmente ou totalmente automatizadas controladas, individualmente ou coletivamente, por softwares. Existe hoje todo um mercado de manipulação de interações possibilitado particularmente pelo emprego de bots (Cf. DI RESTA et al., 2017; BRADSHAW e HOWARD, 2019). É possível comprar visualizações no YouTube, votos em enquetes digitais, seguidores no Instagram, Pinterest e Twitter (GU et al., 2017). China, Rússia, Índia e países do Leste Europeu estão dentre as regiões onde frequentemente se hospedam sites de venda desses serviços, que envolvem bots, mas também contas operadas por humanos (DE CRISTOFARO et al., 2014; GU et al., 2017). Um desses serviços, por exemplo, permite que, por meio de uma plataforma web, um cliente possa operar remotamente até dez mil celulares, num tipo de serviço que ficou conhecido como “fazenda de likes” (GU et al., 2017).

A utilização de perfis automatizados distorce o ecossistema de opiniões em ambiente digital criando falsos consensos e ampliando ou reduzindo debates conforme a conveniência de seus operadores. Essa estratégia parece ter um impacto importante no debate político em escala global (BRADSHAW et al., 2020; BRADSHAW e HOWARD, 2019; WOOLEY e HOWARD, 2017). No Brasil, em particular, estudos realizados pela FGV, observaram que robôs do Twitter podem ter sido responsáveis por mais de 11% das interações em momentos da campanha presidencial de 2014 e por cerca de 10% das interações no impeachment presidencial de 2015 (RUEDIGER, 2017; para controvérsias sobre a atuação de robôs, veja-se VOSOUGHI et al., 2018 e BRADSHAW e HOWARD, 2019).

¹⁵ Para as finalidades dessa seção, consideramos em particular os cinco desses elementos: identidade, compartilhamento, relacionamentos, presença e reputação, analisados em ordem respectiva nas seções que trazemos em seguida (Cf. Kietzmann et al., 2011).

Há evidência ainda de que o próprio implusonamento de publicações vendido pelo Facebook incorpore em algum grau curtidas de perfis ilegítimos (DE CRISTOFARO et al., 2014; para estudos de caso, veja-se: FILAN, 2012; CELLAN-JONES, 2012), o que ressalta a incorporação de mecanismos fraudulentos ao negócio das redes sociais.

3.3.2 Ferramentas de compartilhamento

O emprego do termo “viral” no domínio digital vem da observação de que, na internet, a propagação de uma informação pode ser rápida e descontrolada. Mecanismos de compartilhamento, incorporados pela maioria das redes sociais, permitem que conteúdos sejam redistribuídos em cadeia e se tornem virais. Segundo levantamento do Reuters Institute, 58% dos entrevistados disseram utilizar recursos de compartilhamento em meios digitais (NEWMAN et al., 2019). Outro levantamento apontou ainda que 23% dos americanos têm ciência de terem compartilhado notícias falsas (PEW 2016). Desinformação pode ter inclusive, 70% mais chances de ser retuídada na comparação com postagens verdadeiras (VOSOUGHI et al., 2018).

O caso mais exemplar do impacto desse mecanismo na propagação de conteúdo malicioso e desinformação talvez seja o Whatsapp¹⁶. A plataforma de mídia permite atualmente a criação de grupos de usuários com até 256 pessoas e o compartilhamento de mensagens com até 256 contatos pelo recurso “lista de transmissão”. O Whatasapp se tornou um dos principais recursos de compartilhamento de notícias (NEWMAN et al., 2019) e a possibilidade de distribuição em massa de conteúdos parecem ter influenciado fenômenos sociais e políticos no mundo. O compartilhamento de conteúdo malicioso pelo aplicativo está associado, por exemplo, à ocorrência de linchamentos na Índia que chegaram a ficar conhecidos como “linchamentos do Whatsapp” (para uma discussão, veja-se BANAJI et al., 2019 e ARUN, 2019). Acredita-se também que plataforma tenha tido um papel importante no compartilhamento de desinformação durante a campanha presidencial brasileira de 2018, alavancando os votos no candidato eleito (ABDIN, 2019; BENITES, 2018).

De forma semelhante, no Facebook, observou-se que compartilhamentos impulsionaram, às vésperas das eleições de 2018, a formação de redes digitais de desinformação. Uma das mais notáveis entre essas redes estava associada ao site “Política na rede”, que, a partir de 23 páginas articuladas, chegava a atingir 433 milhões de compartilhamentos por semana (ORTELLADO e RIBEIRO, 2018).

Os mecanismos de compartilhamento operacionalizam a transmissão rápida e simples de conteúdos de terceiros, facilitando que materiais sejam propagados de maneira irrefletida. A distribuição através de uma longa cadeia permite, inclusive, que a fonte de determinado conteúdo malicioso possa se manter oculta (SANTOS et al., 2019). Ao mesmo tempo que o compartilhamento é parte fundamental do fenômeno da desinformação é também um elemento orgânico na mecânica da maior parte das grandes redes sociais. Os “trending topics” do Twitter, por exemplo, deixam clara a intenção da plataforma de fomentar a viralidade.

3.3.3 Interconexão de usuários e conteúdos

A interconexão entre usuários e conteúdos é um dos elementos constitutivos das mídias sociais. E a quantidade de pessoas que as mensagens atingem nesses ambientes é um produto da organização em forma de rede desses espaços (SANTOS et al., 2019). Indivíduos e conteúdos funcionam como nós de estruturas policêntricas e interconectadas

¹⁶ O Whatsapp é às vezes tratado não como uma rede social, mas sim como um mensageiro instantâneo. O tratamento do aplicativo como uma rede social aqui segue uma tendência da literatura (eg. RESENDE et al., 2019). Assim como outras redes sociais, o aplicativo permite a formação de comunidades e possibilita a criação de redes, nas quais usuários e conteúdos estão interconectados. O aplicativo, contudo, se distingue de outras redes sociais por alguns aspectos, como a criptografia de ponta-a-ponta, que torna mais difícil o rastreamento de conteúdos falsos. A engenharia do Whatsapp também dificulta que conteúdo malicioso seja tirado do ar, porque as mensagens são armazenadas nos próprios celulares dos usuários e não em servidores.

(CASTELLS, 2004), permitindo que material noticioso seja difundido de maneira desierarquizada, sem a mediação de profissionais da informação. No tocante à desinformação, a difusão viral desses materiais se sedimenta, em parte, nessas redes ocasionalmente orgânicas que se configuram no ambiente das redes sociais.

A interconexão entre usuários e conteúdos é capaz de produzir “ecossistemas” de desinformação, tais quais os que se formaram nas eleições de 2018, interligando sites de notícias, páginas e usuários do Facebook (ORTELLADO e RIBEIRO, 2018). A interconexão também se materializa entre grupos de Whatsapp, sugerindo que mais do que um aplicativo de envio direto de mensagens, a plataforma articula redes de usuários, que podem acabar possibilitando a propagação de desinformação (RESENDE et al., 2019).

Contudo, a interconexão dos conteúdos nas redes sociais nem sempre é espontânea, mas resultado da operação dos algoritmos. Tem havido uma intensa discussão do quanto esses scripts seriam responsáveis pela conformação das chamadas “bolhas” (ou, no inglês, “echo chambers”) de polarização política, entendidas por alguns como um elemento central da desordem informacional (USHER-LAYSER, 2016; TUFEKCI, 2015; DUBOIS e BLANK, 2018). Essas redes homogêneas de usuários se estruturam a partir da exposição dos indivíduos a conteúdo alinhados a sua visão política. Embora as bolhas tenham se tornado mais evidentes no Facebook, há evidência de que o direcionamento de conteúdos seja responsável pela formação de bolhas também em plataformas como o YouTube (BESSI et al., 2016). Um dos argumentos nesse sentido é o de que o YouTube agregaria vídeos que estariam dispersos se não fosse pela ação de seu algoritmo de recomendações. Isso parece ter ajudado, por exemplo, a alavancar audiência para conteúdos de desinformação associados à extrema direita (CÓRDOVA, 2019).

Há evidência também de que redes de desinformação podem ser deliberadamente estruturadas por agentes políticos. Peritos recrutados pelo Supremo Tribunal Federal teriam identificado que campanhas de desinformação seriam articuladas a partir de uma rede de influenciadores financiada por empresários. Conteúdos difamatórios, segundo sugere o ministro relator do processo, seriam distribuídos no Twitter de maneira simultânea e aparentemente coordenada entre os membros dessa rede (STF, 2020). Na campanha presidencial de 2018, houve relatos de que redes de Whatsapp teriam sido montadas por supostos profissionais de marketing digital para impulsionar campanhas. Embora contassem com as interações espontâneas de usuários, as equipes de marketing digital tinham certo controle sobre o funcionamento dessa rede de grupos, permitido, em parte, por um perfilamento dos participantes (FERREIRA e SOARES, 2018).¹⁷

3.3.4 Instantaneidade

O impacto negativo das redes sociais, quando se considera o fenômeno da desinformação atualmente, é uma consequência também da velocidade da propagação de mensagens em ambiente digital. A presença simultânea de emissores e receptores de informação possível em muitas plataformas permite uma comunicação instantânea mesmo à distância. Um conteúdo pode alcançar, assim, quase que imediatamente um número amplo de receptores. Esse é um dos fatores mais importantes na dimensão assumida pela desinformação na contemporaneidade em comparação ao fenômeno da desinformação na era pré-internet. Essa propagação acelerada da desinformação faz ainda com que a reversão de seus efeitos nem sempre seja eficiente.

¹⁷ Note-se que, no caso do Whatsapp, a distribuição dos conteúdos não é mediada por um algoritmo de ranqueamento, como no caso do Facebook, por exemplo. Observou-se também, nesta rede, que grupos organizados por afinidades não políticas, como grupos de família, têm um papel importante na circulação de notícias falsas (ORTELLADO e RIBEIRO, 2018), o que poderia ser um argumento em prol da organicidade da circulação de desinformação no caso do Whatsapp. Contudo, é preciso considerar que estes grupos talvez funcionem mais como receptores do que como difusores dessas mensagens, de modo que a centralidade de grupos organizados por estratégias de marketing digital deve ser levada em conta (SANTOS et al., 2019).

Alguns estudos já documentaram ainda o caráter exponencial da propagação de desinformação nas redes sociais. Num monitoramento de quase 10 mil participantes de grupos de Whatsapp, observou-se que, em menos de 7 horas, uma notícia falsa chegou a quase 70% desses usuários (SANTOS et al., 2019). Outro monitoramento, detectou que o volume de informações falsas sobre o assassinato da vereadora Marielle Franco aumentou 91 vezes num intervalo de apenas 4 dias no Whatsapp (ORTELLADO e RIBEIRO, 2018).

Além disso, talvez por seu caráter apelativo (Cf. seção 4.1), a desinformação tenha sua circulação ampliada quando comparada a outros tipos de conteúdo. Um levantamento sobre imagens de desinformação em grupos do Whatsapp, por exemplo, observou que a propagação desse tipo de material parece ser mais veloz do que a de outros tipos de imagem (Resende et al., 2019). Evidência na mesma direção foi observada em uma análise do Twitter, na qual se verificou que informações classificadas como falsas podiam atingir a marca de 1500 usuários seis vezes mais rápido do que notícias classificadas como verdadeiras (VOSOUGHI et al., 2018).

3.3.5 Influenciadores

A comunicação digital permitiu uma desintermediação relativa do processo de comunicação. É possível a qualquer um produzir conteúdo em formato noticioso nas redes sociais. Esses produtores independentes constroem reputações e podem se tornar referências em determinados domínios.

Tomando o YouTube como exemplo, dos dez canais políticos com maior número de inscritos, metade deles pode ser considerado independente, inclusive os dois maiores — Nando Moura¹⁸ e Mamaefalei¹⁹ (OLIVEIRA et al., 2019)²⁰. Nesse cenário, “influenciadores digitais” e “mídias hiper-partidárias”²¹ (MARWICK e LEWIS, 2017; RIBEIRO e ORTELLADO, 2018) tornam-se proeminentes no cenário político.

Embora isso possa dar a ideia de que as redes promovem pluralidade de opiniões, é preciso notar que a diversidade desse meio é limitada. Ainda no caso do Youtube, numa investigação de 320 canais políticos, observou-se que apenas 1% deles concentra quase 1/4 do total de visualizações. Quando é considerado o número de inscritos, 10% dos canais concentram quase metade do número de assinantes nesse universo (RIBEIRO e ORTELLADO, 2018).

Essa concentração está parcialmente relacionada à crescente monetização das redes sociais. Para que seu material tenha alcance, um produtor de conteúdos precisa ter estratégias definidas quanto à regularidade, à qualidade e ao engajamento de sua produção, o que privilegia aqueles com possibilidades de investimento financeiro e algum grau de profissionalização de sua atuação. No caso do Facebook, por exemplo, o alcance orgânico (ou seja, a exibição pela qual não se paga) de uma postagem é estimado em apenas 2% do número de assinantes de uma página, implicando a necessidade de financiamento para o crescimento nesta plataforma (JOSEPH, 2018).

Além disso, como deixa supor o já mencionado inquérito sobre desinformação do STF, a consolidação de vozes aparentemente alternativas e espontâneas no ambiente das redes pode ser na verdade fruto de uma estratégia pensada e financiada por parte de agentes políticos ou econômicos interessados em determinadas narrativas (STF, 2020). Assim, a emergência de referências e a construção de reputações fora do ambiente da mídia tradicional não implica necessariamente a democratização do universo da comunicação. O funcionamento econômico das redes afasta progressivamente esse ambiente do mito da democracia e da liberdade digital: muitos falam, mas poucos são de fato ouvidos (Cf. HINDMAN, 2018). A configuração das mídias contemporâneas, nesse sentido, não se

¹⁸ <https://www.youtube.com/channel/UCOOCeqi5txwviDZ4M5W9QSg>

¹⁹ <https://www.youtube.com/channel/UcKsJy-IOEq-eMtarZl2uH1Q>

²⁰ Levantamento feito entre agosto e setembro de 2019.

²¹ O conceito se refere a *outlets* de notícias engajados, alinhados estritamente a uma determinada posição política e que produzem material de “combate” para disputas narrativas (Cf. RIBEIRO e ORTELLADO, 2018)

diferencia tanto da configuração das mídias tradicionais, quando se considera que a emergência de vozes dissonantes é sempre permitida, como uma expressão de valores liberais e democráticos, mas também sempre restrita a espaços de menor amplitude (HERMAN e CHOMSKY, 1988).

4. RECEPÇÃO: POR QUE LEMOS, ACREDITAMOS E COMPARTILHAMOS CONTEÚDOS FALSOS?

A produção e financiamento de desinformação só se faz porque há quem consuma, acredite e socialize informações falsas. Parte das pesquisas sobre o fenômeno da desinformação têm buscado, então, compreender como se dá a recepção desse tipo de material. Esta seção apresenta uma revisão de trabalhos com esse objetivo, considerando, em particular, os aspectos psicológicos relacionados à recepção e à socialização de conteúdos falsos. Em primeiro, buscamos analisar o que torna as informações falsas atraentes. Quais elementos e mecanismos permitem que conteúdos falsos capturem a atenção dos indivíduos? Em segundo, é considerada a assimilação dessa informação: buscamos entender, em particular, por que as pessoas podem vir a acreditar nesse tipo de material, até mesmo quando a falsidade é evidente. Por fim, são considerados alguns fatores que levam os indivíduos a socializar conteúdo falso.²²

4.1 A atratividade da desinformação: captura de atenção e estruturação textual

A abundância de informação sobrecarrega nossa atenção, fazendo dela uma commodity disputada, vendida e comprada pelas companhias de tecnologia da informação, no que tem sido chamado de economia da atenção (SIMON, 1971; SHAPIRO e VARIAN, 1999; Cf. MICROSOFT, 2015). Mas como a produção de notícias falsas consegue atrair nossa atenção em um universo de tantas informações potencialmente verdadeiras?

Com a finalidade de terem impacto na opinião pública, divulgadores de conteúdos falsos interpretam as tendências do fluxo de informação no universo digital e lançam mão de estratégias para tornar seu conteúdo mais atraente, numa abordagem que tem sido chamada por alguns de “attention hacking” (BOYD, 2017; MARWICK e LEWIS, 2017).

A organização textual de notícias falsas, por exemplo, é umas das estratégias nesse sentido. Por um lado, as “fake news” mimetizam alguns elementos de textos jornalísticos profissionais, como o layout e a macroestruturação (SOUZA et al., 2018). Por outro lado, alguns elementos apelativos configuram um texto mais atraente para o universo digital. Os títulos de notícias falsas publicadas em sites se diferenciam de notícias reais por serem mais longos, usarem menos “palavras de parada”, menos nomes comuns e mais nomes próprios. Essa configuração poderia facilitar uma leitura que se atenha apenas no título, em consonância com evidências de que o compartilhamento de notícias nas redes nem sempre se traduz em cliques para o acesso a seu conteúdo integral (HORNE e ADALI, 2017; WANG et al., 2016). Em relação ao corpo das notícias falsas, em comparação a notícias reais, este parece ser mais curto, apresentar menos palavras técnicas e mais redundância lexical (HORNE e ADALI, 2017).

Resende e colegas (2019), numa análise de mensagens com conteúdo falso em grupos do Whatsapp, evidenciaram algumas características similares. Essas mensagens são ligeiramente menores e se atêm a um número menor de temas quando comparadas a mensagens não identificadas como falsas. A composição lexical dessas mensagens também sugere o emprego de outras estratégias destinadas a capturar a atenção dos leitores. Termos como “atenção”, “aviso”, “veja”, “olhe”, “ouça” (indicativos de “mensagens de corrente”) são frequentes nas mensagens falsas e parecem aumentar seu potencial de compartilhamento. Palavras de cunho sexual também aparecem com mais frequência em conteúdo falso, em geral em referência a opositores políticos. O uso de exclamações também se mostrou mais

²² Revisões consistentes em língua inglesa sobre a crença e socialização de informações falsas podem também ser encontradas em Forgas e Baumeister (2019) e Greifender et al. (2021).

frequente nessas mensagens, possivelmente numa tentativa de construir algum tipo de apelo emocional (RESENDE et al., 2019), o que condiz com a hipótese de que a excitação (arousal) seja um dos principais gatilhos da viralidade nas redes sociais (BERGER e MILKMAN, 2012).

A análise de publicações no Twitter sugere ainda conteúdos falsos parecem trazer um grau maior de novidade do que notícias verdadeiras, sendo ainda associados a mais reações de surpresa (VOSOUGHI et al., 2018). Informação nova, na medida em que permite atualizar nossas crenças sobre o mundo, pode ser um fator potencialmente atrativo nesses materiais, já que a novidade pode ser fator importante na captura de atenção (ITTI e BALDI, 2009; VOSOUGHI et al., 2018 mas veja-se BERGER e MILKMAN, 2012 para uma avaliação do papel da novidade em comparação com a excitação (arousal)).

4.2 Processamento da desinformação: por que as pessoas acreditam em informação falsa?

A informação falsa, para que tenha impacto, além de atrair o receptor, precisa ser processada como verdadeira por pelo menos uma parte daqueles que a recebem. Pensar o problema da desinformação, assim, passa, necessariamente pelo mapeamento dos fatores que levam os indivíduos a assimilarem conteúdos falsos como verdadeiros. A presente seção apresenta algumas hipóteses e estudos que buscam explicar a recepção de desinformação do ponto de vista psicológico e cognitivo.

4.2.1 Fluência de processamento e familiaridade

Avaliar a falsidade de afirmações não necessariamente requer engajamento cognitivo consciente. Schwarz et al. (2016) revisam algumas evidências sugerindo que processos psicológicos mais sutis podem estar envolvidos nesse processo de avaliação. Fatores associados à facilidade, ou em outros termos, à fluência de processamento (Cf. FAZIO et al., 2015) de informações, defendem os autores, podem levar afirmações a serem mais frequentemente julgadas como verdadeiras. Por exemplo, mesmo qualidades irrelevantes de um ponto de vista racional, como a simplicidade da fonte, o contraste da letra ou a presença de rimas parecem contribuir para que uma afirmação seja mais frequentemente julgada como verdadeira (REBER e SCHWARZ, 1999; MCGLONE e TOFIGBAKHS, 2000).

Nessa direção, há evidência de que a exposição prévia a uma informação falsa, ou seja, a sua familiaridade, torna mais provável ela seja julgada como verdadeira, mesmo para sujeitos que não estão conscientes da repetição (PENNYCOOK et al., 2015). A exposição pode inclusive favorecer o julgamento de conteúdos pouco plausíveis, embora ainda haja controvérsia sobre essa possibilidade em particular (Cf. DECHENE et al., 2010; PENNYCOOK et al., 2015).

4.2.2 Influência social e confiança

Não é coincidência que o impacto de práticas de desinformação tenha ganhado notoriedade com a expansão das redes sociais. Nesses ambientes, a confiança que as pessoas têm em seus contatos sociais pode ser explorada para dar credibilidade à informação, inclusive no caso de conteúdos falsos. Um experimento de 2007, em que se estudou o fenômeno do *phishing*, é um exemplo desse mecanismo. Foi observada a predisposição dos indivíduos a confiar em e-mails falsos supostamente emitidos por seus contatos sociais. Neste estudo, 72% das pessoas que recebiam um e-mail falso de um remetente que tinha o nome de um suposto conhecido forneciam informações de segurança sensíveis, ao passo que apenas 16% dos indivíduos que recebiam esse e-mail de um remetente sem vínculo social forneciam os mesmos dados (JAGATIC et al., 2007).

A credibilidade de notícias também é impactada pelo tipo de relação social que um indivíduo tem com a fonte da mensagem. A confiabilidade da fonte, por exemplo, é um dos fatores que modula a avaliação dos receptores de conteúdo jornalístico em mídias sociais, predispondo-os a se engajar com esse conteúdo (STERRET et al., 2018). Nesse mesmo sentido, participantes de outro estudo declararam ter uma propensão maior de interagir com

notícias no Facebook enviadas por amigos próximos (BUCHANAN e BENSON, 2019) quando comparadas a notícias enviadas por pessoas pouco próximas.

A percepção de presença social (ARGO et al., 2005.) é um dos fatores que diminuem a suspeita dos indivíduos em relação a conteúdos. Em tarefas em que se deveria avaliar a veracidade de afirmações e nas quais se contava com a possibilidade de recorrer a um mecanismo de *fact-checking*, indivíduos se mostraram menos propensos a acessar a checagem quando tinham a informação de que não se encontravam sozinhos, mas sim em um grupo que realizava a mesma tarefa (JUN et al., 2017). O resultado sugere que a mera percepção de presença social pode reduzir a vigilância de indivíduos em relação a afirmações.

Há evidência, inclusive, de que os produtores de conteúdos maliciosos tentam se valer de alguma forma da possível influência social na recepção de informações. Numa análise lexical de mensagens de grupos de Whatsapp, observou-se o uso mais frequente da palavra “nós” em materiais falsos, quando comparados a materiais verdadeiros, o que poderia ter o propósito de agregar a comunidade em torno da ideia ou conteúdo (RESENDE et al., 2019). A utilização de perfis automatizados (Cf. seção 3.3.1) também é um exemplo de como o mercado de desinformação tenta se valer da suscetibilidade dos indivíduos a fatores sociais e, mais particularmente, ao consenso social (Cf. SCHWARZ et. al, 2016).

4.2.3 Personalidade e envolvimento pessoal

Parte dos estudos sobre a recepção de desinformação têm tentado avaliar em que medida determinados indivíduos poderiam estar mais propensos a confiar ou interagir com conteúdos falsos. Nessa perspectiva, em um estudo que avaliou a propensão de interação com notícias no Facebook e sua relação com o modelo dos Cinco Grandes Fatores de personalidade (extroversão, neuroticismo, amabilidade, realização, abertura à experiência – COSTA e McRAE, 1992) e com a propensão ao risco (MEERTENS e LION, 2008), observou-se que o traço amabilidade (agreeableness, também traduzido socialização – Cf. SILVA e NAKANO, 2011) parece ter impacto na propensão à interação: pessoas com maiores scores de amabilidade têm menor propensão à interação com as notícias. Uma vez que o estímulo noticioso apresentado no estudo abordava um tema negativo (corrupção), os autores sugerem que pessoas com maior score nesse traço tenderiam a evitar a interagir com esse tipo de conteúdo em razão da sua tendência amigável de personalidade (BUCHANAN e BENSON, 2019).

O envolvimento pessoal, ou em outros termos o grau de relevância pessoal de uma informação, é um outro exemplo de fator subjetivo que modula a crença e o potencial de interação com conteúdos. Profissionais de saúde, por exemplo, afirmaram estarem mais dispostos a confiar e a compartilhar rumores sobre saúde cujos temas eles haviam ranqueado como importantes em termos pessoais (CHUA e BANERJEE, 2018). O interesse em temas políticos também parece modular o comportamentos de consumo de notícias dos indivíduos (DUBOIS e BLANK, 2018).

4.2.4 Letramento

O grau de inserção de um indivíduo nas práticas sociais de escrita, ou seja, seu grau de letramento (Cf. SOARES, 1998; KLEIMAN, 2005), parece ser também um aspecto psicossocial importante no fenômeno da desinformação. Em particular, o letramento midiático, que envolve habilidades de processar, usar e avaliar informações midiáticas, considerando as habilidades de lidar com as tecnologias da informação (Cf. REGIS et al. 2014; GOMES et al., 2020) pode ser um fator relevante para entender a dimensão assumida pela desinformação no universo digital.

Uma relação entre o grau de letramento e o problema da desinformação é sugerida, por exemplo, pela observação de que indivíduos que relatam consumir mais (e portanto diversos) canais de mídia tendem a ter hábitos que evitam as chamadas bolhas de informação, como ler materiais com as quais não concordam e implementar estratégias de checagem (DUBOIS e BLANK, 2018). A falta de letramento midiático é uma das explicações

apontadas também para resultados indicando que idosos tendem a compartilhar até sete vezes mais notícias de sites associados a notícias falsas do que jovens entre os 18 e 29 anos (GUESS et al., 2019; para uma revisão, veja-se ESTABEL et al., 2020).

Há resultados que indicam ainda que a educação midiática diminui o impacto de conteúdos falsos. Estudantes com maior grau de treinamento midiático tendiam a rotular mais frequentemente postagens baseadas em evidências como verdadeiras do que postagens contendo informações falsas (KAHNE e BOWYER, 2017). No mesmo sentido, há ainda evidência de que informações falsas sobre segurança alimentar são identificadas mais facilmente por pessoas com maior pontuação na escala de auto-eficiência de uso da internet²³ (WANG et al., 2020), um instrumento que mede habilidades relacionadas ao letramento informacional e midiático (Cf. GOMES et al., 2020)

4.2.5 Viés de confirmação e dissonância cognitiva

Os seres humanos tendem a construir crenças e convicções baseadas em suas expectativas e essas expectativas fazem com que a busca e a interpretação de evidências aconteçam para confirmar somente o que se acredita. Esse fenômeno é estudado pela Psicologia desde a segunda metade do século XX e recebeu o nome de viés confirmatório ou enviesamento de confirmação (WASON, 1960; NICKERSON, 1998; MØLLER & SKAANING, 2021). O viés de confirmação ocorre em dois momentos distintos: na escolha da fonte de informação e na opção por privilegiar um grupo de evidências em detrimento de outras igualmente disponíveis, mas que contrariam as convicções a priori do sujeito (TALLURI et al., 2018). O conceito de viés confirmatório comumente se refere a mecanismos não intencionais, embora em algumas circunstâncias a intencionalidade desses processos possa ser difícil de detectar. O'Sullivan e Schofield (2018) sugeriram que o mecanismo mais comum que leva a erros diagnósticos na clínica médica é o enviesamento. O viés de confirmação responderia por 75% de todos os equívocos médicos. O profissional de medicina tende a observar certos sintomas e relatos do paciente e isso o torna predisposto a escolher os exames que privilegiam sua hipótese diagnóstica inicial em detrimento a uma investigação mais completa. Neste mesmo sentido, Møller e Skaaning (2021) sugerem que os estudos históricos de muitos historiadores seguem a mesma lógica, buscam evidências de fontes específicas que tendem a confirmar suas hipóteses iniciais e negligenciam ou diminuem a importância de evidências contrárias que refutariam suas suposições.

Com a explosão das redes sociais e a difusão descontrolada da informação a partir da segunda metade da década de 2010, pesquisadores vêm usando o constructo do viés confirmatório para compreender o posicionamento político da população. Há evidência de que crenças de natureza política poderiam distorcer o processamento de informação com implicações para processos como de memória, julgamento e visão (CARUSO et al., 2009; CASTELLI e CARRARO, 2011; KAHAN et al., 2012; FREIDA et al., 2013; GRANOT et al., 2014; VAN BAVEL e PEREIRA, 2018). Del Vicario e colaboradores (2017) apresentaram resultados interessantes em relação às escolhas das notícias divulgadas em redes sociais e o enviesamento confirmatório usando inteligência artificial e simulação computacional. Nas simulações, as notícias mais polarizadas eram as mais acessadas, com isso, os algoritmos das redes sociais que privilegiam maior engajamento passavam a aumentar a frequência de notícias de um polo ou de outro do espectro valorativo, fazendo com que, em dado momento, somente um tipo de fonte de informação polarizada chegasse ao usuário. Em outras palavras, as redes sociais e as notícias veiculadas através de algoritmos de engajamento (que buscam aumentar a audiência) podem criar um mundo paralelo, uma espécie de bolha de informação que só apresenta evidências que suportam as crenças a priori dos usuários que, por sua vez, perdem a capacidade de se expor ao contraditório (DEL VICARIO et al., 2017).

Existe ainda uma tendência de seleção de informações que apresentam conteúdo negativo sobre o lado oposto do espectro político, portanto, os usuários de redes sociais tendem a preferir notícias que são deletérias à imagem de oponentes políticos ou que exaltam

²³ Do inglês, *internet self-efficacy scale* (KIM e GLASSMAN, 2013).

positivamente as qualidades do polo apoiado, mesmo que sejam notícias exageradas ou falsas em seu conteúdo (WESTERWICK et al., 2017). Na eleição de Donald Trump, Tappin, Van Der Leer e McKay (2017) encontraram evidências de as fontes de informação mais polarizadas eram as que geram mais adesão, com especial atenção ao fato de que as propostas de governo foram aquelas que menos geravam engajamento por parte dos eleitores. Já Knobloch-Westerwick, Mothes e Polavin (2020) foram além, mostrando que notícias falsas exageradas que possuem conteúdos negativos em relação ao polo oposto do espectro político tendem a gerar mais engajamento do que notícias falsas positivas a favor do lado apoiado pelos eleitores e mais do que qualquer tipo de notícia verdadeira. Os resultados dessa pesquisa ainda denunciam o mecanismo de produção de notícias falsas que trazem fatos parciais e enviesados de notícias que são originalmente verdadeiras para dar um tom de veracidade no discurso e servir de fundamento para confirmar o viés do eleitor. Em outras palavras, uma fake news política eficaz para os propósitos de engajamento em redes sociais e plataformas de notícias é aquela que é polarizada, que tende a acusar os adversários políticos com conteúdos negativos, mas que carrega em si algum fato que realmente aconteceu para dar uma roupagem de veracidade.

Outra abordagem parcialmente relacionada também sugere uma tendência de zelar pela consistência do sistema de crenças. A teoria da dissonância cognitiva (FESTINGER, 1957) é uma das perspectivas que também têm sido consideradas no tratamento do fenômeno da desinformação (TADDICKEN e WOLFF, 2020; WANG et al. 2020). Segundo essa teoria, de caráter cognitivo, mas também emocional, conhecimentos conflitantes gerariam um desconforto que os indivíduos tenderiam a evitar (Para uma discussão relacionada, veja-se também o debate acerca do “raciocínio motivado”: Cf. KAHAN, 2013; KAHNE e BOWYER, 2017; SCHAFFNER e ROCHE, 2017; PENNYCOOK e RAND, 2019; OLIVEIRA e CORTÊS, 2019). Para reduzir esse desconforto, os indivíduos descartariam ou reduziriam a importância de cognições dissonantes ou, alternativamente, buscariam ou valorizariam cognições consonantes (para uma revisão e explicações alternativas a essa teoria, veja-se HARMON-JONES e MILLS, 2017). Indivíduos que têm um elevado grau de confiança na segurança dos alimentos, por exemplo, se tornam mais propensos a compartilhar notícias ruins sobre segurança alimentar, quando comparados a indivíduos com menor grau de confiança. À luz da teoria da dissonância cognitiva, essa tendência poderia ser explicada pelo fato de que essas notícias seriam mais aversivas para os indivíduos com alto grau de confiança, de modo que a replicação do conteúdo poderia ser uma estratégia de redução da ansiedade: o compartilhamento aliviaria o desconforto, como uma espécie de “desabafo” (WANG et al. 2020). Há evidência, por outro lado, que conteúdo dissonante com o ponto de vista de um indivíduo pode se tornar mais aceitável após exposição reiterada (PENNYCOOK et al., 2015).

4.2.6 Processamento Heurístico

De maneira geral, os seres humanos estão inclinados a economizar recursos cognitivos. Em certas circunstâncias, julgamentos e tomadas de decisão podem ser levados a cabo de maneira econômica, considerando informações parciais e rapidamente acessíveis. Estratégias dessa natureza, sejam elas implementadas de maneira consciente ou automática, são comumente entendidas como formas de processamento heurístico. Na literatura, em geral, a heurística é tratada como uma alternativa a processos analíticos e racionais, considerados cognitivamente mais onerosos (Cf. FISCHHOFF, 1999; STANOVICH, 1999; para controvérsias em relação à oposição entre heurística e racionalidade, veja-se KAHAN, 2013).

Processos heurísticos parecem ser subjacentes a diferentes domínios da cognição, desde o processamento da linguagem (Cf. KARIMI e FERREIRA, 2016) a processos integrativos de alto nível, como tomadas de decisão. Neste último domínio, por exemplo, Samuel Popkin (1995) sugeriu uma série de “atalhos” que os eleitores considerariam para avaliar partidos e candidatos em disputas eleitorais. Uma vez que procurar, processar e analisar informação são processos que demandam tempo e esforço, um eleitor pode, por exemplo, assumir uma estratégia de baixo custo para escolher seu candidato considerando

fatores como etnia, religião ou sexo, assumindo esses fatores como proxys de suas tendências políticas e preditores de sua atuação futura (POPKIN, 1995). A fidelidade de eleitores a posições partidárias, inclusive, poderia ser considerada uma forma de heurística (BAVEL e PEREIRA, 2018).

Alguns dos fatores psicológicos e cognitivos já elencados nesta seção como pertinentes à recepção de conteúdos falsos também podem ser considerados como formas de processamento heurístico. Nesse sentido, tendências como a de tomar informação mais familiar ou mais fácil de processar como verdadeira (Cf. seção 4.2.1), a de confiar em conteúdos transmitidos por contatos sociais (Cf. seção 4.2.2) ou a de se interpretar evidência nos termos de crenças e expectativas prévias (Cf. seção 4.2.5) podem ser entendidas como estratégias econômicas para o julgamento de conteúdos (para uma revisão relacionada, veja-se também: SWCHARZ et al., 2016).

A observação de que notícias falsas, quando comparadas a notícias reais, apresentariam títulos densos e apelativos e, ao mesmo tempo, um corpo de textos mais redundante e com menos informação técnica poderia indicar ainda que a produção de notícias falsas é de alguma forma direcionada a despertar processos menos analíticos e mais heurísticos. O texto, assim, não parece tentar convencer o leitor por meio de uma argumentação racional (HORNE E ADALI, 2017).

Além da tendência à redução de custo cognitivo, é possível considerar que o desengajamento de processos analíticos e racionais poderia ainda ser potencializado por outros fatores, como, por exemplo, aqueles relacionados a uma diminuição da precaução dos indivíduos em relação a informações. Nesse sentido, há resultados indicando que a presença social poderia tornar os indivíduos menos cautelosos na análise de afirmações (JUN et al., 2017; Cf. seção 4.2.2). Outro fator relevante poderia ser ainda a propensão de considerarmos nós mesmos ou membros de grupos aos quais pertencemos como menos suscetíveis a conteúdos falsos do que terceiros (JANG e KIM, 2018).

4.3 Por que compartilhamos desinformação?

Esta seção discute, em particular fatores de natureza psicossocial que influenciariam particularmente o compartilhamento, ou seja, a socialização de conteúdos e que poderiam estar envolvidos, assim, no potencial de viralização de conteúdos falsos.²⁴

Um dos gatilhos de compartilhamento mais comumente apontados é a valência emocional da informação. Informações negativas parecem despertar um potencial maior de engajamento e compartilhamento. Uma das observações nesse sentido, por exemplo, sugere que rumores pessimistas nos quais produtos comuns eram associados à incidência de câncer despertavam mais intenções de compartilhamento quando comparados a rumores positivos que associavam alimentos a propriedades curativas (CHUA e BANERJEE, 2018). Há também evidência de que a maior atividade de usuários em comunidades de teorias conspiratórias está relacionada com um comportamento emocional mais negativo dessas comunidades (DEL VICARIO et al., 2016). A análise de mensagens de desinformação no Whatsapp, inclusive, detectou que estas costumam ser ligeiramente mais negativas quando comparadas a informação verdadeira (RESENDE et al., 2019). Da mesma forma, comentários e postagens em páginas conspiratórias no Facebook tendem a ser mais negativos que em páginas científicas (ZOLLO et al., 2015).

Segundo uma das hipóteses que explicaria a relação entre a negatividade e a tendência ao compartilhamento, conteúdo negativo teria, em primeiro lugar, o potencial de atrair a atenção das pessoas e, além disso, despertaria picos intensos de emoção negativa. O compartilhamento, segundo essa hipótese, seria uma espécie de resolução para este

²⁴ Uma vez que o compartilhamento é tratado, em muitos estudos, como indicativo de que o receptor acreditou em determinado conteúdo, alguns dos fatores pertinentes à socialização de conteúdo falso já foram abordados na seção anterior, embora a tônica da mesma seja a crença nesse tipo de conteúdo. Assim, já se comentou na seção anterior que a influência social (4.2.2), envolvimento pessoal (4.2.3) e as crenças ou convicções prévias (4.2.5) poderiam predispor o indivíduo a compartilhar informação falsa.

sentimento negativo, despertado em particular quando a informação é dissonante das crenças do indivíduo (WANG et al., 2020; Cf. seção 4.2.5).

Há, no entanto, controvérsias nesse sentido. No caso de notícias verdadeiras, uma análise de artigos do New York Times detectou que conteúdos com valência positiva tendem a ser mais compartilhados do que conteúdo negativo. Levando em conta, entretanto, que artigos associados a sentimentos negativos, como ansiedade e raiva, também aparecem entre os mais compartilhados do jornal, é possível considerar que mais do que a valência, talvez a excitação (arousal), na verdade, seja o fator preponderante como gatilho de compartilhamento (BERGER e MILKMAN, 2012). Não é claro, contudo em que medida esses resultados se repetiriam na análise de conteúdos falsos.

Além de gatilhos de natureza emocional, outros mecanismos são apontados como potencializadores da socialização de conteúdo online. O desejo de afirmação pela sinalização de traços de identidade seria uma outra possível motivação para o compartilhamento de informação. Segundo essa hipótese, compartilhamentos seriam motivados pela vontade de um indivíduo de influenciar a maneira como é visto pelos outros. Assim, uma postagem na internet por exemplo poderia ter a mesma função do que a escolha por determinadas grifes. A possibilidade de aprofundamento de laços sociais poderia, ainda, ser outro fator relevante para o compartilhamento de conteúdo (Cf. MILKMAN e BERGER, 2014).

5.0 CONSIDERAÇÕES FINAIS: QUAL O FUTURO DAS FAKE NEWS?

O impacto da desordem informacional é fruto de uma equação que envolve os potenciais da tecnologia associados a um conjunto de fatores políticos, econômicos e sociais que favorecem a produção, circulação e assimilação de conteúdos falsos e maliciosos.

No que toca ao desenvolvimento tecnológico, o aprendizado de máquina e a inteligência artificial podem se tornar catalisadores da desordem informacional. Um dos exemplos disso são as ferramentas de geração automática de textos, que alimentadas por um banco de dados, podem recorrer à internet para produzir textos coerentes sobre qualquer assunto, alegadamente similares aos produzidos por humanos (Cf. ZELLERS et al, 2019; FLORIDI e CHIRIATTI, 2020). Esse tipo de instrumento, à disposição do mercado da desinformação, pode ter um potencial perigoso se incorporado à propaganda computacional. De maneira similar, passam a se popularizar as tecnologias de produção de conteúdos audiovisuais chamados de *deepfake*. Essa tecnologia, que já foi usada em campanhas políticas com fins legítimos (Cf. CHRISTOPHER, 2020), atualmente é popular no Brasil como meio de sátira e entretenimento (Cf. RUDNITZKI, 2020; FONSECA, 2020). No entanto, seu uso como ferramenta de desinformação, principalmente em espaços nem tão públicos da internet, como os grupos de Whatsapp, tem também um enorme potencial ofensivo em termos políticos e sociais.

No que diz respeito ao campo social e político, há certamente um avanço no debate da desordem informacional, que tem se expressado, dentre outras coisas, pela proliferação de trabalhos acadêmicos e pelo tratamento recorrente do tema na mídia e até no cinema. Embora, no exterior, haja avanços tanto na regulação pública quanto em iniciativas de caráter educacional, no Brasil, o debate tem produzido resultados ainda tímidos. As iniciativas legislativas patinam nos interesses da classe política, que se por um lado se indigna quando atingida pela desinformação, por outro lado teme perder o uso potencial desse mercado como forma de propaganda política eficaz. Nessa seara, o surgimento de movimentos como o Qanon, uma metateoria conspiratória com ares messiânicos, não deixa de gerar preocupação (Cf. FANJUL, 2021). Integrantes do movimento estiveram entre os responsáveis pela invasão do Capitólio em 2021, motivados por alegações de fraude eleitoral em prol de uma conspiração internacional. De maneira ainda menos expressiva, o movimento já vem encontrado espaço no Brasil (Cf. REBELO, 2020).

Na inércia dos poderes legislativo e executivo, a ofensiva mais significativa e sistemática contra a desordem informacional tem se dado no plano do judiciário, em particular, no que até o momento vem sendo chamado de “Inquérito da Fake News” (STF, 2020). A investigação relacionada ao processo tem ajudado a deslindar alguns dos mecanismos e

personagens mais importantes da teia da desinformação no país. Contudo, a impessoalidade dessa iniciativa é algo que ainda deverá ser avaliado no seu desenrolar.

REFERÊNCIAS

ABDIN, L. "Bots and fake news: the role of WhatsApp in the 2018 Brazilian Presidential election." *Casey Robertson*, v. 41, p. 1, 2019.

ADORNATO, A. C. "Forces at the gate: Social media's influence on editorial and production decisions in local television newsrooms." *Electronic News*, v. 10, n. 2, p. 87-104, 2016.

ALLCOT, H.; GENTZKOW, M. "Social media and fake news in the 2016 election." *Journal of economic perspectives*, v. 31, n. 2, p. 211-36, 2017.

ARUN, C. "On WhatsApp, Rumours, and Lynchings." *Economic & Political Weekly*, v. 54, n. 6, p. 30-35, 2019.

AVAAZ. "As fake news estão nos deixando doentes?." SOCIEDADE BRASILEIRA DE IMUNIZAÇÕES, SBIM, 2019.

B. MCGLONE, M. S., & TOFIGHBAKSH, J. (2000). Birds of a feather flock conjointly (?): Rhyme as reason in aphorisms. *Psychological Science*, 11, 424–428

BANAJI, Shakuntala et al. WhatsApp vigilantes: an exploration of citizen reception and circulation of WhatsApp misinformation linked to mob violence in India. 2019.

BANCO DO BRASIL. *Banco do Brasil esclarece notícias que repercutem decisão do TCU.*

BANCO DO BRASIL, 2020. Disponível em:

<<https://www.bb.com.br/portalbb/page118,3366,3367,1,0,1,0.bb?codigoNoticia=61492>>.

Acesso em: 02 ago. 2020.

BAYER, Judit et al. "Disinformation and propaganda—impact on the functioning of the rule of law in the EU and its Member States." *HEC Paris Research Paper No. LAW-2019-1341*, 2019.

BAYER, Judit et al. "Disinformation and propaganda—impact on the functioning of the rule of law in the EU and its Member States." *HEC Paris Research Paper No. LAW-2019-1341*, 2019.

BENITES, A. *A máquina de 'fake news' nos grupos a favor de Bolsonaro no WhatsApp.* El País, 28 set de 2018. Disponível em:

<https://brasil.elpais.com/brasil/2018/09/26/politica/1537997311_859341.html>. Acesso em:

05 jul de 2020.

BERGER, Jonah; MILKMAN, Katherine L. "What makes online content viral?." *Journal of marketing research*, v. 49, n. 2, p. 192-205, 2012.

BESSI, Alessandro et al. "Users polarization on Facebook and Youtube." *PloS ONE*, v. 11, n. 8, p. e0159641, 2016.

BOYD, Danah. "Hacking the attention economy." *Data and Society: Points*. Disponível em: <https://points.datasociety.net/hacking-the-attention-economy-9fa1daca7a37>, 2017.

BRADSHAW, S.; BAILEY, H.; HOWARD, P. "Industrialized Disinformation: 2020 Global Inventory of Organized Social Media Manipulation." *Oxford Internet Institute*, 2021.

BRADSHAW, S; HOWARD, P. N. "Challenging truth and trust: A global inventory of organized social media manipulation." *The Computational Propaganda Project*, v. 1, 2018.

BRADSHAW, S; HOWARD, P. N. "The global disinformation order: 2019 global inventory of organised social media manipulation." *Project on Computational Propaganda*, 2019.

BUCHANAN, T; BENSON, V. "Spreading Disinformation on Facebook: Do Trust in Message Source, Risk Propensity, or Personality Affect the Organic Reach of "Fake News"?" *Social Media+ Society*, v. 5, n. 4, p. 2056305119888654, 2019.

CARUSO, E.M. et al. "Political partisanship influences perception of biracial candidates' skin tone." *Proceedings of the National Academy of Sciences USA* 106, 20168–20173, 2019.

CASTELLS, M. "The network society: A cross-cultural perspective." *Northampton, MA*. Edward Elgar, 2004.

CELLAN-JONES, R. *Who 'likes' my Virtual Bagels?*. BBC News, 2012. Disponível em: < <https://www.bbc.com/news/technology-18819338>>. Acesso em: 02 jul 2021.

CHADWICK, ANDREW; HOWARD, PHILIP N. (Ed.). *Routledge handbook of Internet politics*. Taylor & Francis, 2010.

CHOMSKY, N. *Mídia: propaganda política e manipulação*. WMF Martins Fontes, 2013.

CHRISTOPHER, N. *AI-generated fake videos that are notoriously rampant in porn are now infiltrating politics*. VICE, 2020. Disponível em: <<https://www.vice.com/en/article/jgedjb/the-first-use-of-deepfakes-in-indian-election-by-bjp>>. Acesso em: mar.2021.

CHUA, AYK; BANERJEE, S. "Intentions to trust and share online health rumors: An experiment with medical professionals." *Computers in Human Behavior*, v. 87, p. 1-9, 2018.

CÓRDOVA, Y. *Como o YouTube se tornou um celeiro da nova direita radical*. The Intercept Brasil, 2019. Disponível em: < <https://theintercept.com/2019/01/09/youtube-direita/> >. Acesso em: 10 jul. 2020.

COSTA, P. T., & MCCRAE, R. R. "Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO FFI): Professional manual." *Psychological Assessment Resources*, 1992

DE ALMEIDA OLIVEIRA, A. M.; DE REZENDE CÔRTEZ. "Ideologia e propensão à crença em Fake News¹." *Fabrcio Bertini Pasquot Polido Lucas Costa dos Anjos. Políticas, Internet e Sociedade*, p. 142, 2019.

DE CRISTOFARO, E., FRIEDMAN, A., JOURJON, G., KAAFAR, M. A., & SHAFIQ, M. Z. "Paying for likes? understanding facebook like fraud using honeypots." *In Proceedings of the 2014 Conference on Internet Measurement Conference*, pp. 129-136, 2014.

DECHÊNE, A. et al. "The truth about the truth: A meta-analytic review of the truth effect." *Personality and Social Psychology Review*, v. 14, n. 2, p. 238-257, 2010.

DEL VICARIO, M. et al. "Echo chambers: Emotional contagion and group polarization on facebook." *Scientific reports*, v. 6, p. 37825, 2016.

DEL VICARIO, M., SCALA, A., CALDARELLI, G., STANLEY, H. E., & QUATTROCIOCCHI, W. "Modeling confirmation bias and polarization". *Scientific Reports*, 7(1), 1-9, 2017. doi: 10.1038/srep40391

DESAI, S. C.; REIMERS, S. "Some misinformation is more easily countered: An experiment on the continued influence effect." *In: CogSci*, 2018.

DI RESTA, Renee et al. *The Bots That Are Changing Politics*. Vice, Nov, v. 2, p. 2017, 2017. Disponível em: < <https://www.vice.com/en/article/mb37k4/twitter-facebook-google-bots-misinformation-changing-politics> > Acesso em: 02 jul 2021.

DOMINGOS, R. *É #FAKE que Bolsonaro se aposentou do Exército por insanidade mental*. G1, 2018. Disponível em: <<https://g1.globo.com/fato-ou-fake/noticia/2018/11/17/e-fake-que-bolsonaro-se-aposentou-do-exercito-por-insanidade-mental.ghtml>>. Acesso em: 23 set. de 2020.

DOS SANTOS, João Guilherme Bastos et al. WhatsApp, política mobile e desinformação: a hidra nas eleições presidenciais de 2018. *Comunicação & Sociedade*, v. 41, n. 2, p. 307-334, 2019.

DUBOIS, E., BLANK, G. "The echo chamber is overstated: the moderating effect of political interest and diverse media." *Information, communication & society*, v. 21, n. 5, p. 729-745, 2018.

VEIGA, E. *A luta anônima de três brasileiros contra sites de fake news*. DW, 2020. Disponível em: <<https://www.dw.com/pt-br/a-luta-an%C3%B4nima-de-tr%C3%AAs-brasileiros-contra-sites-de-fake-news/a-54153272>>. Acesso em: fev. 2021.

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

ESTABEL, L.B.; LUCE, B. F.; SANTINI, L. A. "Idosos, fake news e letramento informacional." *Revista Brasileira de Biblioteconomia e Documentação*, v. 16, p. 1-15, 2020.

ESTADÃO. *CPI das Fake News identifica 2 milhões de anúncios do governo em canais de 'conteúdo inadequado'*. Estadão, 2020. Disponível em: <<https://outline.com/wVgWfu>>. Acesso em: mar. 2021.

ESTADÃO. *Ministro do TCU suspende anúncios do BB em sites acusados de espalhar fake news*. Estadão, 2020b. Disponível em: <<https://outline.com/6FvB34>>. Acesso em: 02 ago. 2020.

FACEBOOK. *Termos de serviço*. Disponível em : <<https://pt-br.facebook.com/terms>>. Acesso em: 23 jul. 2020.

FANJUL, S. *Teorias conspiratórias do QAnon varrem o mundo e são mais perigosas do que parecem*. El país, 2021. Disponível em: <<https://brasil.elpais.com/internacional/2021-01-12/teorias-conspiratorias-do-qanon-varrem-o-mundo-e-sao-mais-perigosa-do-que-parecem.html>>. Acesso em: mar. 2021.

FARAH, T. *Folha Política, famosa por espalhar fake news, só foi banida hoje do Facebook*. BuzzFeedNews, 2018. Disponível em: <<https://www.buzzfeed.com/br/tatianafarah/folha-politica-famosa-por-espalhar-fake-news-so-foi-banida>>. Acesso em: 17 jul. 2020.

FARKAS, JOHAN; SCHOU, JANNICK. *Post-truth, fake news and democracy: Mapping the politics of falsehood*. Routledge, 2019.

FERREIRA, G.; SOARES, J. P. *Como funciona a máquina de WhatsApp que pode eleger Bolsonaro*. Época, 2018. Disponível em: <<https://epoca.globo.com/como-funciona-maquina-de-whatsapp-que-pode-eleger-bolsonaro-23180627>>. Acesso em: 16 jul. 2020.

FESTINGER, L. *A theory of cognitive dissonance*. Stanford university press, 1957.

Fischhoff, B. "Judgment heuristics." *The MIT encyclopedia of the cognitive sciences*, p. 423-425, 1999.

FLORIDI, L.; CHIRIATTI, M. "GPT-3: Its nature, scope, limits, and consequences." *Minds and Machines*, v. 30, n. 4, p. 681-694, 2020.

FONSECA, D. *Bruno Sartori: deepfakes, política e ameaças*. Trip, 2020. Disponível em: <<https://revistatrip.uol.com.br/trip/bruno-sartori-deepfakes-politica-e-ameacas>>. Acesso em: mar. 2021.

FORGAS, JOSEPH P.; BAUMEISTER, ROY (Ed.). *The Social Psychology of Gullibility: Conspiracy Theories, Fake News and Irrational Beliefs*. Routledge, 2019.

GELFERT, A. "Fake news: A definition." *Informal Logic*, v. 38, n. 1, p. 84-117, 2018.

GHEDIN, R; DIAS, T.; RIBEIRO, P. V. *Fake news a R\$ 25 mil por mês: como o Google treinou e enriqueceu blogueiros antipetistas*. The Intercept Brasil, 2019. Disponível em: < <https://theintercept.com/2019/11/19/fake-news-google-blogueiros-antipetistas/> >. Acesso em: 17 jul. 2020.

GLEICHER, N. *Removendo comportamento inautêntico coordenado*. Facebook, 2020. Disponível em: < <https://about.fb.com/br/news/2020/07/removendo-comportamento-inautentico-coordenado/> >. Acesso em: 02 ago. 2020.

GOMES, S. F.; PENNA, J. C. B. O.; ARROIO. "A. Fake News Científicas: Percepção, Persuasão e Letramento." *Ciência & Educação*, v. 26, e. 20018, 2020.

GRANOT, Y. et al. "Justice is not blind: Visual attention exaggerates effects of group identification on legal punishment." *J. Exp. Psychol. Gen.* 143, 2196, 2014.

GREIFENEDER, RAINER et al. *The psychology of fake news: Accepting, sharing, and correcting misinformation*. 2021.

GU, L.; KROPOTOV, V.; YAROCHKIN, F. "The fake news machine: how propagandists abuse the internet and manipulate the public." *Trend Micro*, v. 5, 2017.

GUESS, A.; NAGLER, J.; TUCKER, J. "Less than you think: Prevalence and predictors of fake news dissemination on Facebook." *Science advances*, v. 5, n. 1, p. eaau4586, 2019.

GUNTHER, R.; BECK, P. A.; NISBET, E. C. "Fake news may have contributed to Trump's 2016 victory." (2018). Disponível em: <<https://www.documentcloud.org/documents/4429952-Fake-News-May-Have-Contributed-to-Trump-s-2016.html>>. Acesso em: 23 set. 2020.

JENKINS, H. "Convergence Culture." *New York: New York University Press*, 2006.

HARMON-JONES, E.; MILLS, J. "An introduction to cognitive dissonance theory and an overview of current perspectives on the theory." *Cognitive dissonance: Reexamining a pivotal theory in psychology*, p. 3-24, 2019.

HERMAN, E. S.; CHOMSKY, N. *Manufacturing consent: The political economy of the mass media*. Random House, 1988.

HINDMAN, M. *The myth of digital democracy*. Princeton University Press, 2008.

HITCHEN, J. et al. "Whatsapp and Nigeria's 2019 Elections: Mobilising the People, Protecting the Vote." *AfricaPortal*, 2019.

HORNE, B. D.; ADALI, S. "This just in: Fake news packs a lot in title, uses simpler, repetitive content in text body, more similar to satire than real news." Disponível em: <arXiv preprint arXiv:1703.09398>, 2017.

INSTITUTO DE TECNOLOGIA SOCIAL (ITS). Nota Técnica sobre os Projetos de Lei nº 2927/2020 (Câmara) e nº 2630/2020 (Senado). ITS, 2020. Disponível em: < <https://itsrio.org/pt/publicacoes/nota-tecnica-its-pls-contra-fake-news/> >. Acesso em: fev. 2021.

ITTI, L.; BALDI, P. "Bayesian surprise attracts human attention." *Vision research*, v. 49, n. 10, p. 1295-1306, 2009.

- FILAN, J. *Facebook Ads: What Are You Really Paying For?*. Search Engine Journal, jul.12. Disponível em: < <http://www.searchenginejournal.com/facebook-ads-what-are-you-really-paying-for/46194/>>, Acesso em: Jul. 2012.
- JACK, C. "Lexicon of lies: Terms for problematic information." *Data & Society*, v. 3, p. 22, 2017.
- JAGATIC, T. N. et al. "Social phishing." *Communications of the ACM*, v. 50, n. 10, p. 94-100, 2007.
- JOSEPH, S. 'Organic reach on Facebook is dead': Advertisers expect price hikes after Facebook's feed purge. Digiday, 2018. Disponível em: <<https://digiday.com/marketing/organic-reach-facebook-dead-advertisers-will-spend-reach-facebooks-feed-purge/>>. Acesso em: 30 jul. 2020.
- HABGOOD-COOTE, Joshua (2019) "Stop talking about fake news!" *Inquiry*, 62:9-10, 1033-1065, 2019. DOI: 10.1080/0020174X.2018.1508363
- JUN, Y.; MENG, R.; JOHAR, G. V. "Perceived social presence reduces fact-checking." *Proceedings of the National Academy of Sciences*, v. 114, n. 23, p. 5976-5981, 2017. Argo JJ, Dahl DW, Manchanda RV (2005) The influence of a mere social presence in a retail context. *J Consum Res* 32:207–212.
- K. THOMAS, D. MCCOY, C. Grier, A. KOLCZ, and V. PAXSON. "Trafficking Fraudulent Accounts: The Role of the Underground Market in Twitter Spam and Abuse." *In USENIX Security Symposium*, pp. 195-210, 2013.
- KAHAN, D.M. et al. "They saw a protest: Cognitive illiberalism and the speech-conduct distinction." *Stan Rev* 64, 851 67, 2012.
- KAHAN, Dan M. "Ideology, motivated reasoning, and cognitive reflection: An experimental study." *Judgment and Decision making*, v. 8, p. 407-24, 2013.
- KAHNE, J.; BOWYER, B. "Educating for democracy in a partisan age: Confronting the challenges of motivated reasoning and misinformation." *American Educational Research Journal*, v. 54, n. 1, p. 3-34, 2017.
- KAMINSKA, I. *A lesson in fake news from the info-wars of ancient Rome*. Financial Times, 2018. Disponível em: <<https://www.ft.com/content/aaf2bb08-dca2-11e6-86ac-f253db7791c6>>. Acesso em: 23 set. 2020.
- KARIMI, H.; FERREIRA, F. "Good-enough linguistic representations and online cognitive equilibrium in language processing." *Quarterly journal of experimental psychology*, v. 69, n. 5, p. 1013-1040, 2016.
- KIM, Y., & GLASSMAN, M. "Beyond search and communication: Development and validation of the Internet Self-Efficacy Scale (ISS)." *Computers in Human Behavior*, 29, 1421–1429, 2013
- KNOBLOCH-WESTERWICK, S., MOTHES, C., & POLAVIN, N. Confirmation bias, ingroup bias, and negativity bias in selective exposure to political information. *Communication Research*, 47(1), 104-124, 2020.
- KLEIMAN, A. B. *Preciso "ensinar" o letramento. Não basta ensinar a ler e a escrever*, p. 2005-2010, 2005.
- LAZER, D. MJ et al. "The science of fake news." *Science*, v. 359, n. 6380, p. 1094-1096, 2018.
- LEES, C. "Fake news: the global silencer: The term has become a useful weapon in the dictator's toolkit against the media. Just look at the Philippines." *Index on Censorship*, v. 47, n. 1, p. 88-91, 2018.

LEWANDOWSKY, S. et al. "Misinformation and its correction: Continued influence and successful debiasing." *Psychological science in the public interest*, v. 13, n. 3, p. 106-131, 2012.

LEWIS, R.; MARWICK, A. E. "Taking the red pill: Ideological motivations for spreading online disinformation. Understanding and addressing the disinformation ecosystem." *In Annenberg School for Communication*, p. 18-22, 2017.

LOCATELLI, P. *Partido do pré-candidato general Mourão financiou empresa que mantém site de fake news Folha Política*. The Intercept Brasil, 2018. Disponível em: < <https://theintercept.com/2018/07/11/folha-politica-fake-news-prtb-boatos/> >. Acesso em: 18 jul. 2020.

MAGALHÃES, M. "Fake news" já influenciavam eleições brasileiras em 1945. The Intercept Brasil, 2018. Disponível em: < <https://theintercept.com/2018/02/07/fake-news-ja-influenciava-eleicoes-brasileiras-desde-1945/> >. Acesso em: mar. 2021.

MARWICK, A.; LEWIS, R. "Media manipulation and disinformation online." *New York: Data & Society Research Institute*, 2017.

MEERTENS, R. M., & LION, R. "Measuring an individual's tendency to take risks: The Risk Propensity Scale." *Journal of Applied Social Psychology*, 38(6), 1506–1520, 2008. <https://doi.org/10.1111/j.1559-1816.2008.00357>.

MICROSOFT CANADA. Attention Spans: Consumer Insights, 2015

MILKMAN, K. L.; BERGER, J. "The science of sharing and the sharing of science." *Proceedings of the National Academy of Sciences*, v. 111, n. Supplement 4, p. 13642-13649, 2014.

MOELLER, S. "Weapons of mass destruction and the media: anatomy of a failure." *YaleGlobal Online*, v. 14, 2004.

MØLLER, J., & SKAANING, S-E. "The Ulysses Principle: A Criterial Framework for Reducing Bias When Enlisting the Work of Historians". *Sociological Methods & Research*, 50(1), 103-134, 2021. doi: 10.1177/0049124118769107

MOSCHELLA, M.; WATTS, R. *UK election: what we learned working with Full Fact*. First Draft, 2017. Disponível em: < <https://firstdraftnews.org/latest/joint-venture-learnings/> >. Acesso em: 23 set. 2020.

NEMER, D. *Eduardo e Flávio Bolsonaro são os criadores de grupos de WhatsApp de mentiras contra jornalista da Folha*. The Intercept Brasil, 2020. Disponível em: < <https://theintercept.com/2020/02/14/eduardo-flavio-bolsonaro-criadores-whatsapp-mentiras-jornalista/> >. Acesso em: 02 ago. 2020.

NEWMAN, Nic et al. "Reuters Institute digital news report 2019." *Reuters Institute for the Study of Journalism*, 2019.

OBSERVATÓRIO DA IMPRENSA. "Boni lembra o debate Collor vs. Lula". *Observatório da Imprensa*, 2011. Disponível em < <http://www.observatoriodaimprensa.com.br/videos/boni-relembra-o-lt-br-gt-debate-collor-vs-lula/> >. Acesso em mai 2021.

OLIVEIRA, R. H. F; RIBEIRO, M. M.; ORTELLADO, P. "Uma descrição dos canais políticos do Youtube." *Monitor do Debate Político no Meio Digital*, 2019.

ORTELLADO, P.; RIBEIRO, MM. "O que são e como lidar com as notícias falsas." *Sur. Revista Internacional de Direitos Humanos*, v. 27, p. 5, 2018.

ORTELLADO, P.; RIBEIRO, MM.. "A difusão dos boatos sobre Marielle Franco: do WhatsApp aos sites de notícias." *Monitor do debate político no meio digital*, 2018.

ORTELLADO, P.; MORETTO M.R. *Levantamento inédito revela sites e páginas no Facebook que podem influenciar a eleição com Fake News*. Época, 201. Disponível em: <<https://epoca.globo.com/levantamento-inedito-revela-sites-paginas-no-facebook-que-podem-influenciar-eleicao-com-fake-news-23005004>>.

O’SULLIVAN, E., & SCHOFIELD, S. “Cognitive bias in clinical medicine”. *Journal of the Royal College of Physicians of Edinburgh*, 48(3), 225-231, 2018.

PALMER, E. “British Government Bans the Phrase 'Fake News'.” *Newsweek*, 2018. Disponível em: <<https://www.newsweek.com/british-government-bans-phrase-fake-news-1182784>>. Acesso em: 23 set. 2020.

PAPACHARISSI, Zizi. *A private sphere: Democracy in a digital age*. Polity, 2010.

PECK, B. *Fake North Korean news report shows World Cup triumphs over Japan, U.S. and China*. Yahoo!Sports, 2014. Disponível em: <<https://sports.yahoo.com/blogs/soccer-dirty-tackle/fake-north-korean-news-report-shows-world-cup-triumphs-over-japan--u-s--and-china-203939872.html>>. Acesso em: 23 set. 2020.

PENNYCOOK, G.; CANNON, T. D.; RAND, D. G. Prior exposure increases perceived accuracy of fake news. *Journal of experimental psychology: general*, v. 147, n. 12, p. 1865, 2018. FAZIO, Lisa K. et al. Knowledge does not protect against illusory truth. *Journal of Experimental Psychology: General*, v. 144, n. 5, p. 993, 2015.

PENNYCOOK, G.; RAND, DAVID G. “Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning.” *Cognition*, v. 188, p. 39-50, 2019.

POPKIN, S. L. *Information shortcuts and the reasoning voter. Information, participation and choice: An economic theory of democracy in perspective*, University of Michigan Press, p. 17-35, 1995.

POSETTI, J.; MATTHEWS, A. “A short guide to the history of ‘fake news’ and disinformation.” *International Center for Journalists*, v. 7, p. 2018-07, 2018.

REBELO, A. ‘QAnon brasileiro’ segue firme nas redes e se mostra alinhado a movimento de teorias conspiratórias dos EUA. *El país*, 2021. Disponível em: <<https://brasil.elpais.com/brasil/2021-02-13/qanon-brasileiro-segue-firme-nas-redes-e-se-mostra-alinhado-a-movimento-de-teorias-conspiratorias-dos-eua.html>>. Acesso em: mar. 2021.

REBER, A. R., & SCHWARZ, N. (1999). “Effects of perceptual fluency on judgments of truth.” *Consciousness and Cognition*, 8, 338–342.

REGIS, F.; AUDI, G.; MAIA, A. “Do letramento ao letramento midiático: práticas e competências cognitivas na cibercultura.” In: XXXVII Congresso Brasileiro de Ciências da Comunicação, Foz do Iguaçu/PR, 2014.

RESENDE, G. et al. “(Mis) Information Dissemination in WhatsApp: Gathering, Analyzing and Countermeasures.” In: *The World Wide Web Conference*, p. 818-828, 2019.

RIBEIRO, M. M.; ORTELLADO, P. “O que são e como lidar com as notícias falsas: dos sites de notícias falsas às mídias hiper-partidárias.” *Sur- Revista Internacional de Direitos Humanos*, São Paulo, v.15, n.27, p. 71-83, 2018.

RUDNITZKI, E. *Yes, nós temos deepfake*. Pública, 2020. Disponível em: <<https://apublica.org/2020/08/yes-nos-temos-deepfake-brasileiros-sao-o-2o-maior-publico-de-aplicativo-que-troca-rostos-de-politicos-e-celebridades/>>. Acesso em: mar. 2021.

RUEDIGER, M. A. "Robôs, redes sociais e política no Brasil: estudo sobre interferências ilegítimas no debate público na web, riscos à democracia e processo eleitoral de 2018." FGV DAPP, 2017.

SAKAMOTO, LEONARDO. *O que aprendi sendo xingado na internet*. Leya, 2016.

SATARIANO, A.; ALBA, D. *Burning cell towers, out of baseless fear they spread the virus*. The New York Times, 2020. Disponível em: <<https://www.nytimes.com/2020/04/10/technology/coronavirus-5g-uk.html>>. Acesso em: 23 set. 2020.

SCHAFFNER, B. F.; ROCHE, C. "Misinformation and motivated reasoning: Responses to economic news in a politicized environment." *Public Opinion Quarterly*, v. 81, n. 1, p. 86-110, 2016.

SCHWARTZ, J. *Trump gives out 'Fake News Awards' to CNN*. NY Times, Wash Post. Politico, January, v. 17, 2018. Disponível em: <<https://www.politico.com/story/2018/01/17/trump-fake-news-awards-345482>>. Acesso em: 02 jul. 2021.

SCHWARZ, J. *Twelve Years Later, US Media Still Can't Get Iraqi WMD Story Right*. The Intercept. Disponível em: <<https://theintercept.com/2015/04/10/twelve-years-later-u-s-media-still-cant-get-iraqi-wmd-story-right/>>. Acesso em: nov. 2020.

SCHWARZ, N. et al. "Metacognitive experiences and the intricacies of setting people straight: Implications for debiasing and public information campaigns." *Advances in experimental social psychology*, v. 39, p. 127-161, 2007.

SHAPIRO, C. & VARIAN, H. R. *Information Rules: A Strategic Guide to the Network Economy*. Harvard Business Press, 1999.

SILVERMAN, C. et al. *Inside the partisan fight for your news feed*. BuzzFeedNews. 2017. Disponível em: <<https://www.buzzfeednews.com/article/craigsilverman/inside-the-partisan-fight-for-your-news-feed>>. Acesso em: 23 set. 2020.

SILVERMAN, C. *How Teens In The Balkans Are Duping Trump Supporters With Fake News*. BuzzFeed News, 2016. Disponível em: <<https://www.buzzfeednews.com/article/craigsilverman/how-macedonia-became-a-global-hub-for-pro-trump-misinfo>>. Acesso em: 02 ago. 2020.

SIMON, H. A. "Designing organizations for an information-rich world." In: Greenberger M., ed. (ed.) *Computers, communications and the public interest*, vol. 72, 37-52 (Johns Hopkins Press, Baltimore, 1971).

SOARES, M. *Letramento: um tema em três gêneros*. Belo Horizonte: Autêntica, 1998.

SOLL, J. *The long and brutal history of fake news*. Politico Magazine, v. 18, n. 12, p. 2016, 2016. Disponível em: <<https://www.politico.com/magazine/story/2016/12/fake-news-history-long-violent-214535>>. Acesso em: 23 set. 2020.

SONG, H.; SCHWARZ, N. "Fluency and the detection of misleading questions: Low processing fluency attenuates the Moses illusion." *Social Cognition*, v. 26, n. 6, p. 791-799, 2008.

SOUZA, K. A. C. et al. "Características jornalísticas nos sites de Fake News: Uma análise das notícias falsas no caso "La Bête"." *Artista nu no MAM São Paulo*. INTERCOM, 2018

STANLEY, S. *Misinformation and hate speech in Myanmar*. First Draft, 2017. Disponível em: <<https://firstdraftnews.org/latest/misinformation-myanmar/>>. Acesso em: 23 set. 2020.

STANLEY, S. *Misinformation and hate speech in Myanmar*. First Draft, 2017. Disponível em: <<https://firstdraftnews.org/latest/misinformation-myanmar/>>. Acesso em: 23 set. 2020.

STANOVICH, K. E. *Who is rational?: Studies of individual differences in reasoning*. Psychology Press, 1999.

STARBIRD, K. "Disinformation's spread: bots, trolls and all of us." *Nature*, v. 571, n. 7766, p. 449-450, 2019.

STARBIRD, K. "Disinformation's spread: bots, trolls and all of us." *Nature*, v. 571, n. 7766, p. 449-450, 2019.

STENCEL, M.; LUTHER, J. *Fact-checking count tops 300 for the first time*. Duke Reporter's Lab, 2020. Disponível em: < <https://reporterslab.org/fact-checking-count-tops-300-for-the-first-time/>>. Acesso em: mar. de 2021.

STERRET, D. et al. "Who shared it?: How Americans decide what news to trust on social media." *Americanpress Institute*, 2018.

STF. *Mandato de 27 de maio de 2020 relativo ao inquérito 4.781*. 2020. Disponível em: < <http://www.stf.jus.br/arquivo/cms/noticiaNoticiaStf/anexo/mandado27maio.pdf> >. Acesso em: 10 jul. 2020.

SUBRAMANIAN, S. *Inside the Macedonian Fake-News Complex*. *Wired*, 2017. Disponível em: < <https://www.wired.com/2017/02/veles-macedonia-fake-news/> > Acesso em: 02 ago. 2020.

TADDICKEN, M.; WOLFF, L. "Fake News' in Science Communication: Emotions and Strategies of Coping with Dissonance Online." *Media and Communication*, v. 8, n. 1, p. 206-217, 2020.

TALLURI, B. C., URAI, A. E., TSETSOS, K., USHER, M., & DONNER, T. H. "Confirmation bias through selective overweighting of choice-consistent evidence". *Current Biology*, 28(19), 3128-3135, 2018.

TANDOC JR, E. C.; LIM, Z. W.; LING, R. "Defining "fake news" A typology of scholarly definitions." *Digital journalism*, v. 6, n. 2, p. 137-153, 2018.

TAPPIN, B. M., VAN DER LEER, L., & MCKAY, R. T. The Heart Trumps The head: Desirability bias in political belief revision. *Journal of Experimental Psychology: General*, 146(8), 1143, 2017.

TAYLOR, E. et al. "Follow the Money: How the Online Advertising Ecosystem Funds COVID-19 Junk News and Disinformation." *Oxford Internet Institute*, 2020.

TOGOH, I. Google And Amazon Are Inadvertently Funding Covid Conspiracy Sites To The Tune Of \$25 Million. *Forbes*, 2020. Disponível em: < <https://www.forbes.com/sites/isabeltogoh/2020/07/08/google-and-amazon-are-inadvertently-funding-covid-conspiracy-sites-to-the-tune-of-25-million/#15f01be17d96> >. Acesso em: 18 jul. 2020.

TOLEDO, D. *Dono de sites criticados por 'fake news' recebe dinheiro de deputado*. *UOL Eleições*, 2018. Disponível em: < <https://noticias.uol.com.br/politica/eleicoes/2018/noticias/2018/05/26/dono-de-sites-criticados-por-fake-news-recebe-dinheiro-de-deputado.htm?cmpid=copiaecola> >. Acesso em: 18 jul. 2020.

TSE – Tribunal Superior Eleitoral. *Nova campanha do TSE contra a desinformação vai ao ar a partir desta terça-feira (1º)*. TSE, 2020. Disponível em: < <https://www.tse.jus.br/imprensa/noticias-tse/2020/Setembro/nova-campanha-do-tse-contr-a-desinformacao-vai-ao-ar-a-partir-desta-terca-feira-1o> >. Acesso em: mar. 2021.

TUFEKCI, Z. "Facebook said its algorithms do help form echo chambers, and the tech press missed it." *New Perspectives Quarterly*, v. 32, n. 3, p. 9-12, 2015.

TWITTER. *Política de Falsa Identidade*. Disponível em: <<https://help.twitter.com/pt/rules-and-policies/twitter-impersonation-policy>>. Acesso em: 23 jul. 2020.

USHER-LAYSER, N. "Newsfeed: Facebook, filtering and news consumption." Phi Kappa Phi Forum, *Gale Academic OneFile*, vol. 96, no. 3, 2016, p. 18+. Disponível em: https://link.gale.com/apps/doc/A468772963/AONE?u=puc_rjbr&sid=AONE&xid=e458d38b. Acesso em: 11 Jul. 2020.

VALOR ECONÔMICO. *Petrobras diz que veiculou 2 milhões de anúncios em sites de fake news por erro*. Valor Econômico, 2020. Disponível em: < <https://outline.com/ux3Xqz> >. Acesso em: 02 ago. 2020.

VAN BAVEL, J. J.; PEREIRA, A. "The partisan brain: An identity-based model of political belief." *Trends in cognitive sciences*, v. 22, n. 3, p. 213-224, 2018.

VEIGA, E. Contra fake news: brasileiros anônimos fazem sites perderem R\$ 448 mil ao mês. UOL, 2020. Disponível em: < <https://www.uol.com.br/tilt/ultimas-noticias/deutschewelle/2020/07/13/contra-fake-news-brasileiros-anonimos-fazem-sites-perderem-r-448-mil-ao-mes.htm> >. Acesso em: 29 jul. 2020.

VISCARDI, J. M. "Fake News, verdade e mentira sob a ótica de Jair Bolsonaro no twitter." *Trab. linguist. apl., Campinas*, v. 59, n. 2, p. 1134-1157, Aug. 2020. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-18132020000201134&lng=en&nrm=iso>. Acesso em: 23 Mar. 2021. Epub Sep 16, 2020. <https://doi.org/10.1590/01031813715891620200520>.

VOSOUGHI, S.; ROY, Deb; ARAL, S. "The spread of true and false news online." *Science*, v. 359, n. 6380, p. 1146-1151, 2018.

WANG, L. X.; RAMACHANDRAN, A.; CHAINTREAU, A. "Measuring click and share dynamics on social media: a reproducible and validated approach." In: Tenth International AAAI Conference on Web and Social Media, 2016.

WANG, R. et al. "Fake news or bad news? Toward an emotion-driven cognitive dissonance model of misinformation diffusion." *Asian Journal of Communication*, p. 1-26, 2020.

WARDLE, C. "First draft's essential guide to understanding information disorder." *First draft*, 2019.

WARDLE, C.; DERAKHSHAN, H. "Information disorder: Toward an interdisciplinary framework for research and policy making." *Council of Europe report*, v. 27, 2017b.

WARDLE, Claire; DERAKHSHAN, Hossein. Thinking about 'information disorder': formats of misinformation, disinformation, and mal-information. Ireton, Cherilyn; Posetti, Julie. Journalism, 'fake news' & disinformation. Paris: Unesco, p. 43-54, 2018.

WASON, P. C. "On the failure to eliminate hypotheses in a conceptual task". *Quarterly Journal of Experimental Psychology*, 12 (3): 129-40, 1960. doi: 10.1080/17470216008416717

WESTERWICK, A., JOHNSON, B. K., & KNOBLOCH-WESTERWICK, S. Confirmation biases in selective exposure to political online information: Source bias vs. content bias. *Communication Monographs*, 84(3), 343-364, 2017. doi: 10.1080/03637751.2016.1272761

WOOLEY, S. C.; HOWARD, P. N. Computational propaganda worldwide: Executive summary. Computational Propaganda Research Project. *Working Paper*, 2017.

ZELLERS, R. et al. "Defending against neural fake news." *arXiv preprint arXiv:1905.12616*, 2019.

ZOLLO, F. et al. "Emotional dynamics in the age of misinformation." *PloS one*, v. 10, n. 9, p. e0138740, 2015.

DECLARAÇÃO DE CONTRIBUIÇÃO DOS AUTORES

Renê Forster – concepção, planejamento e escrita do manuscrito

Rodrigo Monteiro de Carvalho – escrita e revisão do manuscrito

Alberto Filgueiras – escrita e revisão do manuscrito

Emanuelle Avila – escrita e revisão do manuscrito

DECLARAÇÃO DE CONFLITO DE INTERESSE

Os autores deste manuscrito declaram que não haver conflito de interesse de ordem financeira, comercial, política, acadêmica ou pessoal em relação a este trabalho.

Este preprint foi submetido sob as seguintes condições:

- Os autores declaram que estão cientes que são os únicos responsáveis pelo conteúdo do preprint e que o depósito no SciELO Preprints não significa nenhum compromisso de parte do SciELO, exceto sua preservação e disseminação.
- Os autores declaram que os necessários Termos de Consentimento Livre e Esclarecido de participantes ou pacientes na pesquisa foram obtidos e estão descritos no manuscrito, quando aplicável.
- Os autores declaram que a elaboração do manuscrito seguiu as normas éticas de comunicação científica.
- Os autores declaram que os dados, aplicativos e outros conteúdos subjacentes ao manuscrito estão referenciados.
- O manuscrito depositado está no formato PDF.
- Os autores declaram que a pesquisa que deu origem ao manuscrito seguiu as boas práticas éticas e que as necessárias aprovações de comitês de ética de pesquisa, quando aplicável, estão descritas no manuscrito.
- Os autores concordam que caso o manuscrito venha a ser aceito e postado no servidor SciELO Preprints, a retirada do mesmo se dará mediante retratação.
- Os autores concordam que o manuscrito aprovado será disponibilizado sob licença [Creative Commons CC-BY](#).
- O autor submissor declara que as contribuições de todos os autores e declaração de conflito de interesses estão incluídas de maneira explícita e em seções específicas do manuscrito.
- Os autores declaram que o manuscrito não foi depositado e/ou disponibilizado previamente em outro servidor de preprints ou publicado em um periódico.
- Caso o manuscrito esteja em processo de avaliação ou sendo preparado para publicação mas ainda não publicado por um periódico, os autores declaram que receberam autorização do periódico para realizar este depósito.
- O autor submissor declara que todos os autores do manuscrito concordam com a submissão ao SciELO Preprints.