

PÓS-VERDADE E ASPECTOS EPISTÊMICOS DAS MÍDIAS SOCIAIS: ANÁLISES, CRÍTICAS E PROPOSIÇÕES*

POST-TRUTH AND EPISTEMIC ASPECTS OF SOCIAL MEDIA:
ANALYSIS, CRITICS AND PROPOSITIONS

Bismarck Bório de Medeiros**

RESUMO

Este artigo tem por objetivo analisar relações do fenômeno denominado de pós-verdade e a interação de agentes epistêmicos com a grande quantidade de desinformação disseminada nas mídias sociais. Considerando o contexto histórico-social e as pesquisas voltadas a entender a propagação de desinformação e manipulação de opinião pública, interações entre usuários e os vieses cognitivos envolvidos no processo, será observado o status epistêmico do conteúdo informativo nas mídias bem como de seus agentes, avaliando seus modos de formação de crenças e a confiabilidade entre os integrantes dos denominados grupos epistemicamente perniciosos nessas mídias. Consequentemente, a partir dessa investigação, traremos críticas construtivas às abordagens contidas na literatura e duas soluções: uma dentro de políticas educacionais por meio de disciplinas voltadas ao entendimento crítico e engajado das tecnologias da informação; outra na área da Retórica, com a implementação do que denominamos de *graus de persuasão* de um agente epistêmico com relação a outro, seu grau de separação e do modo de apresentação do conteúdo informativo envolvido em seu argumento. Esta contribuição pode ter impacto prático direto no solapamento dessa estrutura de propagação de desinformação, impacto teórico relativo a aspectos dialógico-argumentativos, e em áreas como Epistemologia Social.

PALAVRAS-CHAVE: argumentação; câmara de eco; desinformação; epistemologia social; pós-verdade.

ABSTRACT

This article aims to analyze relationships between the phenomenon called post-truth and the interaction of epistemic agents with the large amount of disinformation disseminated on social media. Considering the historical-social context and research to understand the progression of disinformation and manipulation of public opinion, the interactions between users and the cognitive lives involved in the process, the epistemic status of informational content in the media will be presented, as well as of its agents, evaluating their ways of belief formation and reliability among members of epistemically pernicious groups in these media. Consequently, based on this investigation, we will make constructive criticisms of the approaches contained in the literature and propose two solutions: one within educational policies through disciplines aimed at the critical and engaged understanding of information technologies; other in the Rhetoric area, with the implementation of what we call *degrees of persuasion* of one epistemic agent in relation to another, their degree of separation and the mode of presentation of informative content involved in their argument. This contribution can have a direct practical impact on undermining this structure of spreading misinformation, theoretical impact on the dialogical-argumentative aspects and in areas such as Social Epistemology.

KEYWORDS: argumentation; disinformation; echo chambers; post-truth; social epistemology.

* Artigo recebido em 18/11/2023 e aprovado para publicação em 20/12/2023.

** Doutorando em Filosofia com ênfase em Lógica Matemática na Universidade Federal de Santa Maria. Pesquisa nas áreas de Filosofia Teórica, Metafilosofia e História das Ciências Naturais e Sociais. E-mail: bismarckborio@gmail.com.

INTRODUÇÃO

Décadas atrás, possivelmente poucos leitores teriam conhecimento sobre este artigo, ou nem ao menos estaríamos falando acerca destes tópicos, se não fosse o rápido desenvolvimento de dispositivos eletrônicos como celulares, *tablets* ou o próprio computador, que contam com *softwares* de acesso às redes de internet. O livre acesso, armazenamento e produção de informações proporcionados por essa revolução informacional, bem como suas consequências epistêmicas, sociais e políticas, mudaram a nossa forma de interagir no mundo e nossas formas de pensar sobre cognição; a confiabilidade da circulação de informações em organizações sociais e até como a formação de opinião ocorre no debate em espaço considerado público – para nosso benefício ou malefício como sociedade. É desses temas que trata este artigo.

No início do século, um assunto que estava recebendo atenção por parte de pesquisadores da área de biblioteconomia e ciências de informação (sigla LIS, em inglês para *Library and Information Sciences*) era a mudança de paradigma gerada pela Internet sobre como se estabelecia a credibilidade das fontes de informação que eram veiculadas em ambiente virtual. Havia a consideração de que gradativamente a determinação dessa credibilidade por alguma autoridade competente com certo poder ou domínio de determinado tópico estava sendo substituída – devido ao advento da Internet 2.0 que proporciona a produção e divulgação de conteúdo pelo usuário por meio de plataformas – à noção de confiabilidade (*reliability*), considerada uma noção mais robusta da qual o conceito de autoridade seria dependente. Lankes (2008, p. 680, tradução nossa) dá uma descrição do termo, comparando-o com a visão por autoridade:

[A] Confiabilidade geralmente se refere a algo ou alguém que é percebido como fidedigno e consistente em qualidade. [...] A confiabilidade é simplesmente a consistência dos dados. Se uma coisa (por exemplo, um fato, substância ou fenômeno) está sujeita ao mesmo tratamento (por exemplo, perguntas, experimentos ou aplicações), produz o mesmo resultado? Se uma abordagem de autoridade for tão simples quando acreditar que um determinado âncora de notícias dará uma resposta credível, então mudar de estação de notícias para estação de notícias procurando por pontos em comum na mesma história é uma abordagem de confiabilidade.

Iniciavam-se projetos com sugestões de categorias e diretrizes de credibilidade informacional, bem como relatado o início de pesquisas sobre avaliação de credibilidade e identificação de autoridade cognitiva sobre informações fornecidas por alguém em plataformas, sites ou mídias sociais, como destacado, por exemplo, em Fogg (2002, 2003) e

Rieh (2010). Obviamente, já havia uma preocupação com a divulgação de informações falsas, porém o otimismo acerca da cultura de transparência e da eficácia de meios para verificação de credibilidade era latente, e acabava por conduzir a previsões igualmente otimistas: o sucesso da Wikipédia^{1 2} e na Internet 2.0 como um todo os impulsionava. Todavia, Lankes explicita ao final de seu artigo “[...] o subjacente paradoxo no qual os usuários são simultaneamente mais responsáveis pelas decisões sobre informação, enquanto tornam-se também mais dependentes dos fornecedores dessas informações e das ferramentas usadas para manipular essas informações” (Lankes, 2008, p. 681-682, tradução nossa), paradoxo este que pode minar a confiabilidade epistêmica da informação e do qual pessoas se aproveitaram da maneira como a internet se estruturou para criar uma enxurrada de desinformação, compondo parte do problema que vem a ser chamado de pós-verdade, como veremos adiante.

O termo pós-verdade tem seu uso relatado pela primeira vez em um artigo na revista *The Nation* pelo dramaturgo Steve Tesich sobre as mentiras institucionalizadas, e sem qualquer consequência jurídica, cometidas pelos governos norte-americanos – fato que ele denominava de *Síndrome de Watergate*. Comparando a supressão da verdade em regimes ditatoriais da supressão governamental norte-americana, ele declara:

Até agora, todos os ditadores tiveram de trabalhar arduamente para suprimir a verdade. Nós, pelas nossas ações, estamos dizendo que isso não é mais necessário, que adquirimos um mecanismo espiritual que pode desnudar a verdade de qualquer significado. De uma forma muito fundamental, nós, como povo livre, decidimos livremente que queremos viver em algum mundo pós-verdade (Tesich, 1992. p. 13, tradução nossa).

Inspirado por Tesich, em 2004 o fenômeno da pós-verdade foi abordado historicamente e culturalmente, partindo do fenômeno contemporâneo da desconfiança de que mais mentiras estão sendo contadas socialmente – com a pós-verdade sendo a prática desonesta de dissimulação e apresentação de razões para manipular a verdade – no livro de Ralph Keyes (2018), *A era da pós-verdade*. Todavia, o termo recebeu atenção pública quando foi citado pelo economista Paul Krugman em sua coluna no *The New York Times* no fim de 2011, ao descrever a campanha do candidato republicano Mitt Romney à presidência dos EUA,

¹ Na época, entre 2005 e 2007, a Wikipédia foi submetida a várias avaliações e pesquisas (ver o próprio verbete Confiabilidade da Wikipédia em https://pt.wikipedia.org/wiki/Confiabilidade_da_Wikipédia), sendo verificado que seus índices de precisão eram semelhantes aos da *Encyclopaedia Britannica* (Giles, 2005). Porém, na contraposição dessa construção de plataformas confiáveis de acesso à informação, começavam a existir críticos a tal tendência de mudança na credibilidade da informação, como o escritor Andrew Keen em seu livro *O culto do amador*, por exemplo.

² Para informações sobre o processo de verificação coletiva e de como se mantém a garantia de qualidade e a confiabilidade dos verbetes da Wikipédia, ver Sundin (2011).

referindo-se a ela como fraudulenta e passível de ser caracterizada dentro do que chamou de política da pós-verdade (*post-truth politics*). Sua real relevância no debate público veio após ser eleita a palavra do ano pelo Dicionário Oxford, em 2016. Seu significado foi definido como “relativo ou referente a circunstâncias nas quais fatos objetivos são menos influentes na opinião pública do que emoções e crenças pessoais”. Porém, veremos que esse conceito tem um contexto historicamente mais amplo e anterior a sua própria formulação e parece englobar mais fatores do que o destacado acima pelo dicionário, direcionando nossa discussão para o ambiente das redes sociais, em que tal fenômeno se tornou mais influente, abrangente e quantificável.

Assim, temos a intenção de basear-nos em análises interdisciplinares e pesquisas dentro da sociologia das redes e ciências cognitivas para entendermos como se estrutura epistemicamente as mídias sociais e embasarmos propostas para melhoria em suas interações que solapam a proliferação em massa de notícias falsas, desinformação e opiniões infundadas em detrimento de informações verdadeiras. Para isso, podemos entender como rastrear essa noção ainda ampla de pós-verdade como um fenômeno anterior à sua conceitualização.

1 ANTEVISÕES E APROXIMAÇÕES CONCEITUAIS DA PÓS-VERDADE

A visão de dissolução da verdade pelos valores, interesses e humores dos indivíduos que compõem, interagem e são mediados por mídias e instituições de determinado meio social não parece ser algo novo. Buscar identificar e compreender os componentes desse fenômeno historicamente é importante para o localizarmos por si só. Um dos principais é a mentira como instrumento de manipulação da opinião³. A filósofa política Hannah Arendt dedicou parte do seu pensamento a discutir a relação entre instituições políticas, debate público e a verdade. No livro *Entre o passado e o futuro* (Arendt, 2011) de 1968, a filósofa alemã busca por que a verdade não seria em princípio um valor intrinsecamente político, demonstrando que filosoficamente a verdade era considerada antagônica da mera opinião – esta última, sim, um pré-requisito à atividade política e podendo ser motivada por desejos e interesses. Arendt

³ Já na década de 20, o intelectual e jornalista Walter Lippmann cunhou o termo *fabricação de consenso* para designar o gerenciamento da opinião pública por órgãos de propaganda governamentais. Sua crítica estava atrelada a como órgãos e mídia da época discutiam sobre o fato de Estados com governos democráticos terem menos poder de coerção e deveriam ter meios mais finos de controle da população – referida por jornalistas como um “rebanho desnortado” – e as consequências morais disso para noções como verdade e transparência quando esses instrumentos eram mal empregados. Edward Herman e Noam Chomsky (1988) retomam e aprofundam melhor o cerne dessa crítica, contrastando o que chamam Modelo de propaganda de alinhamento da mídia com o interesse das elites e do governo com a ideia jeffersoniana de mídia jornalística.

traz a questão de que determinados grupos vêm tratando verdades factuais como se fossem opiniões e, mesmo que a verdade factual não seja tão forte quanto a verdade filosófica e dê conteúdo à opinião – ou como Arendt diz, informa a opinião (Arendt, 2011, p. 237) –, ela pode ser mantida separada da opinião; uma não se reduz à outra.

Dessa forma, são observadas certas propriedades do que é factual, como a de não ser necessário nem autoevidente. Porém, uma verdade factual, para Arendt, não seria fruto de consenso e aceitação, já a opinião, sim, havendo a identificação do real antagonista à verdade: a falsidade deliberada. A mentira, diferente da verdade factual, demanda uma atividade que por vezes propositadamente encontra-se disfarçada de proferimento de opinião, trazendo o problema da fabricação de “fatos” voltados aos seus interesses e proveitos, podendo modular sua mentira a seu bel-prazer para soar até mais convincente que a própria verdade (Arendt, 2011, p. 249-250). Nesse ponto, Arendt faz a distinção dentro da política entre a mentira tradicional diplomática, quando se visa falsificar parte da realidade para ludibriar o inimigo, e a mentira moderna que rearranja todo o contexto factual, com potencial de autoengano no processo. Esta última, para Arendt, manifesta-se em regimes totalitários a plena vista. As consequências dessa prática institucional são social e epistemicamente perniciosas:

[...] a longo prazo, o resultado mais certo da lavagem cerebral é uma curiosa espécie de cinismo – uma absoluta recusa a acreditar na verdade de qualquer coisa, por mais bem estabelecida que ela possa ser. Em outras palavras, o resultado de uma substituição coerente e total da verdade dos fatos por mentiras não é passarem estas a ser aceitas como verdade, e a verdade ser difamada como mentira, porém um processo de destruição do sentido mediante o qual nos orientamos no mundo real – incluindo-se entre os meios mentais para esse fim a categoria de oposição entre verdade e falsidade (Arendt, 2011, p. 255).

Já seu ensaio *Lying in politics* (Arendt, 1972) de 1971, feito devido ao vazamento de documentos secretos do governo norte-americano – caso conhecido como *Pentagon Papers* – a autora se propõe a discutir sobre a mentira, agora como ferramenta, com ênfase em Estados considerados democráticos. Ela retorna à condição de contingência do fato, bem como à agência e à capacidade de alteração dos fatos de quem é versado na arte da mentira, dando destaque a duas variedades novas de mentirosos, que são “[...] os aparentemente inócuos *diretores de relações públicas* no governo que aprenderam seu ofício com a inventividade da *Madison Avenue*”⁴ (Arendt, 1972, p. 7-8, tradução e grifos nossos) e homens que seriam

⁴ Não à toa Arendt destaca essa avenida de Nova York para relacioná-la à função de relações públicas. É onde se concentram os publicitários que tiveram como precursor Edward Bernays, considerado o pai das relações públicas contemporâneas. Ele se baseia na psicologia freudiana – ele era sobrinho de Freud e tinha real interesse

solucionadores de problemas (*problem solvers*) da política externa estadunidense – dos quais as relações públicas também fazem parte.

Estes últimos são relevantes ao contexto dos *Pentagon Papers*, pois são tecnocratas de ação extremamente autoconfiantes; estrategistas realizando, a todo custo, a fabricação de fatos para obter a realidade que desejam, processo que Arendt chama de *defactualização*: manipulação da realidade voltada a cooptar a mente da população, seja para “proteger o presidente” ou “evitar humilhação governamental”. Enganando e se autoenganando no processo, visavam obter aceitação, adesão e consenso⁵ ou dissenso (dependendo do tópico), perdendo contato com o mundo real:

[...] os solucionadores de problemas que conheciam todos os fatos que lhes são regularmente apresentados nos relatórios da comunidade de inteligência tiveram apenas de confiar nas suas técnicas partilhadas, isto é, nas várias formas de traduzir qualidades e conteúdos em quantidades e números para com eles calcular resultados – que, inexplicavelmente, nunca se concretizaram – para eliminar, dia após dia, o que sabiam ser real. [...] os solucionadores de problemas não julgaram; eles calcularam. Sua autoconfiança nem mesmo precisava de autoengano para ser sustentada em meio a tantos julgamentos errados, pois se apoiava na evidência da verdade matemática, puramente racional. Exceto, é claro, que esta “verdade” é totalmente irrelevante para o “problema” em questão. (Arendt, 1972, p. 36-37, tradução nossa).

Com esse misto de burocracia técnica, autoengano e defactualização, eles justamente manipularam fatos e a opinião pública, porque foi efetivo e “[...] o desrespeito pela realidade era inerente à política e aos próprios objetivos” (Arendt, 1972, p. 42, tradução nossa). Ou seja, a verdade e os fatos não importavam. Esses relatos de Arendt são um belo exemplo do que viria a ser chamado atualmente de pós-verdade⁶. O termo é recente e tem uma ambiguidade característica, como mostra a linguista Lúcia Santaella (2018, p. 36):

Para o Dicionário, por sua vez, a “pós-verdade” deve ser entendida em dois sentidos diferentes: de um lado, o significado “depois que a verdade tenha se tornado conhecida”, de outro lado, o significado inaugurado pelo artigo de Tesich, a saber, o

pelo trabalho do tio – e as noções da famigerada *Propaganda* (seu primeiro livro levou este título, porém após a Segunda Guerra o termo ficou mal visto, fazendo com que Bernays renomeasse sua expertise). Ele buscou elaborar técnicas baseadas na teoria das pulsões e do inconsciente para moldar e manipular opiniões e desejos das massas para expansão de mercado, gerando maior aceitabilidade ou desejabilidade do consumo de determinado produto, denominando sua função de *engenharia do consentimento*. O documentário da BBC *The century of the self* mostra detalhes interessantes sobre o tema. Para uma introdução acadêmica mais extensa, ver Samuel (2010).

⁵ Aqui podemos perceber melhor que basicamente a defactualização de Arendt é parente da fabricação de consenso de Lippmann e Chomsky.

⁶ Outro exemplo, este já de uma definição proximal, pode ser lido no ensaio de Harry Frankfurt *Sobre falar Merda (On Bullshit)*, de 1986: para o autor, falar merda (*bullshitting*) não seria necessariamente mentir, mas *falar buscando persuadir sem qualquer consideração para com a verdade*.

fato de que a verdade se tornou irrelevante. Assim, no seu sentido expandido, o prefixo “pós” não mais significa apenas “depois de um evento ou situação específica” como, por exemplo, na expressão “pós-guerra”, mas também implica “um tempo em que um conceito se tornou irrelevante ou sem importância”, como foi o caso de pós-nacional.

Nesse contexto, podemos tomar como tese e observar a noção norteadora de pós-verdade como essa desvalorização da verdade e manipulação da factualidade em um primeiro plano, tendo como consequência uma maior relevância pessoal e social no que acreditamos e como nos sentimos acerca disso; conjunto de valores esses derivados historicamente da formação de opinião promovida estrategicamente por instituições econômicas, políticas e midiáticas⁷. (Bennett; Livingston, 2021), por exemplo, elaboram a origem desse emaranhado institucionalizado de desinformação a partir da necessidade de gestão e controle das massas pela “engenharia do consenso” proporcionado pelas relações públicas, visando à estabilidade de governos democráticos. As novas tecnologias e plataformas de mídias sociais ampliaram e tornaram as técnicas de formação de opinião mais apuradas e acessíveis de serem direcionadas a grupos sociais menores e mais selecionados, seja acidentalmente nas bolhas epistêmicas (*epistemic bubbles*) ou intencionalmente, como ocorre nas câmaras de eco (*echo chambers*), fenômenos sociais de que trataremos adiante.

Tal como o filósofo Lee McIntyre (2018, p. 11-13) destaca em seu livro *Post-Truth*, tão mais importante seja estabelecer o que é a pós-verdade seria entender o porquê de estar ocorrendo. Ele compassa sua elaboração de resposta na promoção do negacionismo científico por certas corporações, seja pela indústria do tabaco para desacreditar pesquisas sobre os danos do fumo à saúde, seja por companhias de petróleo para desenvolver contraposições ao fato científico de que o fenômeno do aquecimento global é causado em parte pela queima de combustíveis fósseis. Inclusive, financiaram pesquisas alternativas para produzir informações falsas ou tendenciosas e investiram em *marketing* e *lobby* político para gerar controvérsias e dissenso nas mídias acerca da pesquisa científica, influenciando a opinião pública e a mobilização contrária à criação de políticas e medidas governamentais para mitigar ou restringir os efeitos do consumo de cigarros e da mudança do clima, concluindo que “[...] o uso seletivo de fatos que sustentam uma posição e a rejeição completa de fatos que não sustentam parece parte integrante da criação de uma nova realidade de pós-verdade”

⁷ Historicamente as atividades de manipulação da opinião pública sempre tiveram ligação intrínseca à propagação de notícias falsas. Dutilh Novaes e Ridder (2021) argumentam que o fenômeno considerado contemporâneo das notícias falsas (*fake news*) não é algo recente, sendo identificado em diferentes épocas e modelos distintos de manipulação de opinião, fazendo uso das tecnologias de informação disponíveis e, portanto, com diferentes tipos de implementação, somente.

(McIntyre, 2018, p. 32, tradução nossa). McIntyre então delinea o que ele considera os outros quatro fatores que culminaram nesta era de pós-verdade:

- a) A existência de vieses cognitivos: das pesquisas em Psicologia tendo como paradigmas a teoria da dissonância cognitiva de Leon Festinger e na teoria da conformidade social de Solomon Asch, até as mais contemporâneas que têm como principais vieses o *backfire effect* e o efeito Dunning-Krueger, apontam como somos suscetíveis a falhas e tendenciosos em muitas das maneiras de raciocinarmos para formarmos crenças e realizarmos julgamentos;
- b) Declínio da mídia tradicional: o direcionamento a um partidarismo midiático, incentivo a entrevistas e debates de opinião – com uma alegada imparcialidade de mídia (valor falso ou ineficientemente empregado, claro) – incutindo opiniões perniciosas no debate público (colocando um cientista e um terraplanista para debaterem, por exemplo) e o advento das mídias sociais enfraqueceram a mídia jornalística *mainstream*;
- c) Ascensão das redes sociais e do espalhamento de notícias falsas (*fake news*): a facilidade de interação, disseminação e acesso à informação contrastada com uma maior dificuldade de rastreabilidade e critérios de confiabilidade da fonte dessa informação geram ambientes virtuais epistemicamente vulneráveis à propagação de notícias falsas, desinformação e à criação do que McIntyre chama de cilo informacional, ou bolhas;
- d) a crítica pós-modernista à noção de verdade objetiva: a elaboração de teorias construtivistas sociais da ciência que punham em xeque a concepção de objetividade influenciou setores conspiracionistas, fundamentalistas e políticos de extrema direita a se utilizarem desta crítica para justificar suas crenças, compromissos ideológicos e ações (McIntyre, 2018, p. 32, tradução nossa).

Suas conexões e *insights* são interessantes, porém há críticas que podemos trazer e certos esclarecimentos a essas atribuições. Começamos admitindo que a afirmação de (d) é um bocado duvidosa, pois mesmo que o próprio Bruno Latour (2004) tenha feito um *mea culpa*, a alegação de influência pós-moderna pressupõe ao menos um entendimento da crítica do que e como a objetividade está “em jogo” (como o argumento de que os valores de estruturas de poder têm relevância nas formas de produção de conhecimento, por exemplo), portanto exige leitura e elaboração teórica com base nas literaturas do tema. Jargões e citações de efeito embasando opiniões pragmáticas e enviesadas sobre as relações entre poder, verdade e conhecimento podem ser consideradas, no máximo, tendo uma vaga inspiração pós-modernista. Inspiração e influência são relações distintas. McIntyre (2018, p. 140, tradução nossa) diz que “[...] mesmo se políticos de direita e negacionistas não tiverem lido Derrida e Foucault, a semente da ideia aflorou neles”: como se pós-modernistas tivessem criado tal ideia. Já vimos em Arendt que a concepção de relativização da verdade já estava socialmente implícita em bases filosóficas e culturais distintas do pós-modernismo⁸. O que nos faz pensar:

⁸ Em Bennett e Livingston (2018, p. 12, tradução nossa) eles citam uma fala de Ivy L. Lee, um dos precursores das relações públicas do início do século passado, que se encaixa bem neste contexto: “O que é um fato? O

será o pós-modernismo não influenciador, mas um dos influenciados por uma era de pós-verdade que já vinha despontando no horizonte de possibilidades? Esta é uma alegação que nos parece mais plausível, porém escapa do escopo de nosso artigo.

Dessa forma, observemos também que (a) é uma condição cognitiva que já existe previamente a esse fenômeno sociocultural. Vieses cognitivos podem até fazer parte integrante do *background* social de pós-verdade, mas apenas porque há a criação e o incentivo de ambientes epistêmicos à formação de opiniões sem informações confiáveis e apelo a aspectos afetivos atrelados às atitudes doxásticas dos indivíduos de um grupo – tema de que falaremos adiante. Algo semelhante ocorre tanto com o negacionismo científico quanto (b) e (c): há um elaborado processo no qual pessoas ou instituições disseminam intencional e estrategicamente informações falsas ou enganosas, i.e., *a propagação de desinformação* é um problema crucial em uma realidade de pós-verdade. Dessa forma, temos aqui o início do que irá ser o centro de gravidade deste artigo.

Guiando-nos por Feigold (2022), Doowan Lee e Mills Golub – dois experts em análise de desinformação e cofundadores da Veracity Authentication Systems Technology (VAST) – comentam que campanhas de propagação de desinformação se dão de quatro formas: engenharia social, amplificação inautêntica, microdirecionamento e/ou através de assédio ou abuso gerando engajamento. Vamos analisar como ocorrem esses espalhamentos de desinformação e suas consequências epistêmicas dentro das mídias sociais a partir da próxima seção.

2 MÍDIAS SOCIAIS E FORMAÇÃO DE CRENÇAS EM GRUPO

Como esclarecimento inicial, vamos tratar aqui que ambientes epistêmicos (*epistemic environments*) são “[...] constituídos por fatos, sejam eles sociais ou físicos, que influenciam a realização epistêmica (adquirir conhecimento, compreensão e assim por diante) naquele ambiente” (Ryan, 2021 *apud* Anderau, 2023, p. 107, tradução nossa). Um agente epistêmico é qualquer sujeito que participe passiva ou ativamente de algum ambiente epistêmico, podendo realizar certas atividades – sejam elas obter, transmitir ou compartilhar informações – e deter certas atitudes proposicionais ou capacidades como crer/acreditar, conhecer, compreender ou ter opiniões acerca de algo – uma definição também semelhante à de Anderau. Notícias falsas (*fake news*) podem ser caracterizadas como má-informação ou desinformação, dependendo das intenções dos agentes epistêmicos envolvidos em sua propagação, divulgadas em qualquer

esforço para estabelecer um fato absoluto é simplesmente uma tentativa de fornecer-lhe a minha interpretação dos fatos”.

meio de comunicação. Opinião aqui é caracterizada como uma crença – podendo ser motivada por certos interesses e valores – que o agente busca embasar e justificar com informações, todavia ela é aleticamente neutra, podendo outro agente concordar ou discordar dela e eles debaterem sem qualquer compromisso epistêmico cientificamente preciso envolvido e sem necessariamente algum deles estar errado – portanto, falaremos a partir daqui mais em termos de formação de crença que de formação de opinião. E, por fim, podemos nos basear em um complementar e razoável conceito de desinformação que o próprio Facebook detém:

Conteúdo de informação impreciso ou manipulado que é espalhado intencionalmente. Isto pode incluir notícias falsas ou envolver métodos mais sutis, como operações de *false flag*, alimentar citações imprecisas ou histórias a intermediários inocentes ou amplificar conscientemente informações enviesadas ou enganosas. A desinformação é distinta da má-informação, que é a disseminação inadvertida ou não intencional de informações imprecisas sem intenção maliciosa (Weedon; Nuland; Stamos, 2017, p. 5, tradução nossa).

Posto isso, alguns estudos empíricos de impacto podem nos dar um vislumbre de como se estrutura a relação entre as informações e os agentes epistêmicos nesses ambientes virtuais. Investigando pelo Twitter por onze anos cascatas de notícias, foi observado por Vosougui, Roy e Aral (2018) que notícias falsas atingem mais rápido as pessoas do que notícias verdadeiras. Tal dissonância foi explicada por meio de dois aspectos: o fator emocional ser mais presente em notícias falsas que nas verdadeiras, e pelos graus de novidade envolvidos na propagação de desinformação. Outro resultado da pesquisa é a igualdade da taxa de influência entre *bots* sociais e perfis de pessoas reais, i.e., não foram constatadas diferenças de influência artificial maior que a humana no compartilhamento de notícias falsas, o que nos soaria a princípio contraintuitivo. Contudo, em uma pesquisa mais recente direcionada a notícias envolvendo COVID-19 (Himelein-Wachowiak *et al.*, 2020) o número de *bots* sociais era predominante na discussão do tema, muitos compartilhando excessivamente (Ferrara *et al.*, 2016) e alavancando postagens de baixa credibilidade. Dessa forma, com a constatação por Shao *et al* (2018) de que pessoas compartilham em proporção semelhante a *bots* e outras pessoas, a compreensão da participação dos *bots* sociais pode se dar pela criação de engajamento em torno de notícias de baixa credibilidade, e a própria dinâmica entre os usuários nas mídias realiza o resto do esforço.

No complemento do fenômeno acima, há muitas vezes a estratégia de fabricação de notícias online em série – com etapas definidas, como em uma linha de produção. Tal prática foi relatada em uma organização política populista italiana denominada Movimento Cinco Estrelas, descrita no livro *Engenheiros do caos* (Empoli, 2019), que a partir de 2012 passam a

possuir uma rede de informações capaz de produzir uma realidade paralela como divulgação por seus próprios canais da seguinte forma:

Elas já são recortadas, sob medida, para viralizar no Facebook e nas outras redes sociais. Os títulos são sedutores, muitas vezes enganosos, outras vezes violentos. Começam quase sempre com as mesmas palavras e expressões: *Vergonhoso, Péssima notícia, Isto é a Itália!, Vocês vão ficar chocados, Basta!, É o fim!*. De início, antecipa-se a emoção, em geral negativa, que se quer suscitar. Depois, divulgada a informação, às vezes verdadeira, mas muito frequentemente falsa, convida-se à participação: *Compartilhe!, Faça circular, Máxima difusão!* O único critério de seleção, bem entendido, são os cliques. As notícias que suscitam as reações mais intensas são valorizadas, republicadas, aprofundadas. Tornam-se objeto de discursos e de iniciativas políticas, cavalos de batalha do Movimento. Outras, tediosas, mesmo quando mais importantes e exatas, terminam o dia no pano de fundo, dando espaço às denúncias de complô e de corrupção, reais ou imaginárias (Empoli, 2019, p. 38-39).

Contudo, além da própria interação entre *bots* sociais e usuários, existem tipos de estrutura das redes com vieses algorítmicos, podendo direcionar de modo prejudicial a própria forma de dispor o conteúdo da plataforma baseado em como classifica seus usuários, criando maior entrincheiramento⁹ e polarização (Franke, 2022), influenciando na própria formação de crenças dos agentes ali engajados (Peralta *et al*, 2021), e até sendo constatada uma propensão de agentes com mente aberta (*open-mindedness*) – considerada uma virtude intelectual – a serem manipulados por propaganda externa, seja ela moderada ou extremista (Pansanella *et al.*, 2023). Além disso, Ceylan, Anderson e Wood, (2023) apontam que divulgação de má-informação possa estar ligada aos hábitos de compartilhamento adquiridos com o uso frequente das mídias sociais, tornando-se epistemicamente insensíveis ao conteúdo informacional devido à recompensa de relevância e visibilidade que certos algoritmos fornecem a tais usuários engajados, sendo observado que 15% desses que compartilham habitualmente são responsáveis por cerca de 30 a 40% das notícias falsas espalhadas. Além de que indivíduos quando expostos repetidas vezes a informações falsas tendem a crer nelas, mesmo quando a informação contradiz algum conhecimento prévio do agente (Pillai; Fazio, 2021).

Por sua vez, os estudos de Pennycook e Rand, (2021) vão contra o senso comum de que compartilhamento de notícias falsas tem como fator primário influências políticas e forte partidarização, alegando com base na abordagem clássica de duplo processamento da psicologia cognitiva que um agente com uma maior capacidade de deliberação reduz a probabilidade de obter crenças em notícias falsas, bem como seu discernimento reduzido – devido a não recorrer

⁹ Termo utilizado ao apego por parte do agente epistêmico a crenças firmemente estabelecidas, tornando-se difícil ou indesejável sua revisão ou mudança, mesmo confrontado com evidências contrárias.

ou à falta de conhecimento prévio – faz com que tal agente falhe em distinguir notícias falsas de verdadeiras, mostrando outras pesquisas que demonstravam uma participação forte de comportamentos cognitivamente enviesados (denominados também de heurísticos) nas tomadas de decisão, como de familiaridade e ancoragem, sendo constatada também certa dissociação entre a crença no conteúdo e seu compartilhamento, i.e, nem sempre o que é compartilhado está alinhado com aquilo em que se crê. Tal comportamento pode ter como explicação justamente os resultados de Ceylan, Anderson e Wood, (2023), por exemplo. Assim, no artigo sobre dinâmicas de formação de opinião (Moussaïd *et al.*, 2013), foi constatado por meio de simulações que há dois fatores que influenciam fortemente a opinião dos integrantes de um grupo, denominados de *atratores de opinião*:

1. uma quantidade de pessoas que compartilham opiniões similares, chamado de efeito majoritário (*majority effect*);
2. a existência no grupo de uma ou poucas pessoas com alta credibilidade ou confiáveis, chamado de efeito do especialista (*expert effect*).

Ambos os atratores tendem a puxar os outros integrantes do grupo em direção às suas próprias opiniões, podendo não ser benéficos aos agentes epistêmicos envolvidos e afastá-los de formarem opiniões verdadeiras, além de esses agentes epistêmicos incidirem em dois tipos de vieses: o efeito *bandwagon* e viés de expertise.

Logo, vemos a partir dessas investigações como a estrutura de espalhamento de desinformação nas redes funciona e como os seus usuários, considerados aqui como agentes epistêmicos, afetam e são afetados por tal dinâmica. Tais análises serão decisivas para pensarmos em formas de como contornar a disseminação de desinformação nas redes. Contudo, ainda precisamos abordar como esses temas se conectam com outros resultados e análises sobre estrutura social das redes, ciências cognitivas e epistemologia social, em que a estrutura das mídias sociais, agentes epistêmicos mal intencionados e mal informados, vieses cognitivos e confiabilidade ou credibilidade entre agentes criam ambientes epistemicamente perniciosos.

3 CÂMARAS DE ECO E BOLHAS EPISTÊMICAS: CONFIABILIDADE EM MÍDIAS SOCIAIS

Já expomos características e relações importantes das mídias sociais com o fenômeno da pós-verdade e como ocorre a interação nelas com informações verdadeiras ou falsas por seus usuários. No entanto, há mais camadas a serem exploradas que envolvem um maior entrelaçamento entre o ambiente epistêmico em que o agente se encontra e seus integrantes.

Ora, mesmo se buscarmos mitigar de forma ingênua a proliferação de notícias falsas, encontraremos novos problemas. Temos como exemplo em Thorson (2015) o fenômeno cognitivo denominado de “ecos de crença” (*belief echoes*): mesmo que uma pessoa seja exposta à correção de uma má-informação anterior que baseava certas crenças e exercia efeito sobre as atitudes nessa pessoa, essas atitudes ainda são observadas, havendo a perseverança da crença (*belief perseverance*). Interpretada como uma variação do viés cognitivo chamado de efeito tiro pela culatra (*backfire effect*), sua ocorrência já foi descrita em mídias sociais envolvendo informações falsas corrigidas sobre a COVID-19 e vacinação (Sharevski *et al.*, 2022). Portanto, vemos que a correção de informações não tem uma efetividade tão garantida diretamente. Além do que, mesmo não havendo uma relação estrita entre o uso de mídias sociais e crenças em más informações, quanto mais a pessoa for exposta a elas, mais há correlação entre a exposição do agente e sua formação de crença nessas informações falsas ou conspiratórias (Enders *et al.*, 2021).

Dessa forma, chegamos aos dois principais fenômenos sociais de redes que se relacionam entre si e nos quais convergem as consequências de praticamente todos os estudos que citamos anteriormente: as bolhas epistêmicas (*epistemic bubbles*) e as câmaras de eco (*echo chambers*). Uma bolha epistêmica¹⁰ é uma estrutura epistêmica social em que se adquire informação por meio de uma variedade de recursos limitados, excluindo, por omissão em seu ambiente, agentes epistêmicos relevantes, tendo sua constituição por processos orgânicos de seletividade social e formação de comunidade. Já as câmaras de eco têm boa parte dessas características, mais o reforço de determinadas crenças e a deliberada exclusão de outros tipos de agentes que possam ameaçar a predominância de tais crenças reforçadas e aumentar a coesão do grupo. As crenças mantidas nesses grupos em ambos os casos são fortes entre seus integrantes, porém de baixa confiabilidade contrastada com informações relevantes acerca do conteúdo da crença desses agentes, tornando esse conjunto de agentes epistêmicos *grupos epistemicamente perniciosos* (Boyd, 2018), como veremos adiante.

Nas bolhas epistêmicas, devido à omissão de agentes epistêmicos relevantes e com isso de conteúdos informativos que seriam importantes para se estabelecerem crenças verdadeiras – ou ao menos opiniões bem justificadas – não há boa cobertura por parte dos integrantes que forneça garantias de que as informações que circulam na bolha são confiáveis. Como naturalmente tendemos a nos organizar socialmente em ambientes com pessoas que

¹⁰ As definições aqui traçadas têm como bases os artigos de Nguyen (2020) e Boyd (2018).

tenham algo em comum conosco¹¹ – entre elas concordâncias de opinião fortalecidas pelas próprias mídias que estruturam a experiência de interação entre seus usuários buscando otimizá-la visando ao seu maior tempo de acesso, Lúcia Santaella (2018, p. 13) explicita parte desses mecanismos:

As fontes para a geração de filtros personalizados incluem a história de busca do usuário, o resultado de suas escolhas, sua interação com provedores de serviços, seus interesses demonstrados por produtos e serviços. Além disso, tudo que se posta e compartilha nas redes sociais é também engolido pelos algoritmos de captura do perfil do usuário. [...] a viabilização dos serviços compromete a privacidade e limita a visão de mundo do usuário, estreitando seus horizontes. [...] As bolhas são constituídas por pessoas que possuem a mesma visão de mundo, valores similares e o senso de humor em idêntica sintonia.

Isto é, em bolhas epistêmicas encontra-se maior consonância entre seus integrantes, e com isso uma maior confiança em suas próprias crenças, correndo-se o risco de todos os agentes na bolha reforçarem a autoconfiança nas crenças um do outro; em Boyd (2018) este fenômeno social é chamado de corroboração retroalimentativa (*bootstrapping corroboration*). Esse elemento de retroalimentação aparece no processo de agrupamento (*groupstrapping*): quando um agente epistêmico é membro de um grupo o considerando fonte de testemunho – i.e., que nos fornece condições para constituirmos conhecimento acerca de algo – e à medida que um grupo está garantido a acreditar em algo, está em função das garantias relevantes de seus membros (Boyd, 2018, p. 7), então naturalmente os integrantes do grupo irão dar mais garantias para a crença do grupo nesse algo, recorrendo ao próprio grupo, sem nenhuma evidência externa a ele. Um círculo epistemicamente vicioso e nada virtuoso na formação e reforço de crenças em grupo.

Já uma câmara de eco tem características semelhantes às bolhas epistêmicas, contudo há em sua constituição componentes agencialmente intencionais exclusivos de argumentos contrários a desinformações que muitas vezes baseiam e reforçam crenças predominantes promovidas nesses grupos, bem como a exclusão de agentes epistêmicos relevantes para contraposição a tais informações perniciosas, sendo elas descredibilizadas dentro do próprio ambiente e credibilizando quem traz mais informações (sejam elas informações verdadeiras, porém interpretativamente enviesadas, sejam desinformações) reforçando o sistema de

¹¹ Com as mídias sociais esta visão de “naturalidade” muda de figura, porém os efeitos são os mesmos: a questão nas mídias é o viés algorítmico, que faz uma filtragem personalizada e faz chegar ao usuário recomendações de pessoas e conteúdo apenas que o código reconhece para o agente epistêmico como interessante com base no seu histórico de acesso – e dependendo da rede, essa filtragem tem muito mais influência na criação dessa seletividade. No caso acima, chamamos este fenômeno de filtro de bolha (*filter bubble*). Para mais informações, ler Pariser (2011) e a própria Santaella (2018).

crenças estabelecido no ambiente. Nguyen (2020, p. 46, tradução nossa) traz uma definição com destaque a tais propriedades, sendo a câmara de eco

[...] uma comunidade epistêmica que cria uma disparidade significativa na confiança entre membros e não membros. Tal disparidade é criada pela exclusão de não membros através de descrédito epistêmico, enquanto simultaneamente amplifica as credenciais epistêmicas de seus membros. Finalmente, as câmaras de eco são tais que o acordo geral com algum conjunto básico de crenças é um pré-requisito para a adesão, onde essas crenças centrais incluem crenças que apoiam essa disparidade na confiança. Por “descrédito epistêmico”, quero dizer que os não membros não são simplesmente omitidos ou não ouvidos, mas lhes são ativamente atribuídos algum demérito epistêmico, como falta de confiabilidade, maldade ou desonestidade. Por “amplificar credenciais epistêmicas”, quero dizer que os membros recebem níveis muito altos de confiança.

Há outras definições que são extremamente similares das bolhas epistêmicas, como a que podemos encontrar em Baumgaertner (2014, p. 2549-2550), porém a definição de Nguyen apresenta fatores manipulativos por parte de outros agentes epistêmicos que fundamentam a câmara de eco em uma estrutura que privilegia confiabilidade mútua entre os membros do grupo, e os agentes são direcionados a reforçar discordâncias a certas crenças e afirmações, gerando críticas e rechaçamento a elas baseados no próprio conjunto de crenças do grupo, uma característica moldada intencionalmente e distinta das bolhas epistêmicas. Esse tipo de mecanismo é extremamente nocivo a qualquer possibilidade de revisão de crenças e checagem de informações que sejam epistemicamente virtuosas no sentido dos agentes epistêmicos voltarem-se à verdade e aos fatos.

Assim, vemos que o conceito de pós-verdade nas mídias sociais e a propagação de desinformação estão intrínseca e profundamente atreladas às câmaras de eco. Há pesquisas buscando a análise dessas câmaras de eco em tipos distintos de mídias sociais (Cineli *et al.* 2021), ou em grupos específicos adeptos de teorias conspiratórias (Ruiz; Nilsson, 2022) investigando como essa desinformação circula e ganha tração, reforçando controvérsias direcionadas a identidade de grupo – nós contra eles – em que há retoricamente elementos sociais (*ethos*) e aspectos emocionais (*pathos*) envolvendo reclamações e contendas com base nas crenças de identificação de grupo estabelecidas na câmara de eco¹². Neste trabalho foram reconhecidas duas etapas que constituem a propagação de desinformação nestes ambientes:

¹² Dentro do tema abordado aqui, o capítulo do livro *Defining and defending a insular interpretative community* (Jamieson; Capella, 2008, p. 194-207) relata interessantemente como a opinião da mídia conservadora pela criação de definições compartilhadas, linguagem emocionalmente evocativa – utilização de termos e temas sensíveis ao público – e polarização podem criar um senso de pertencimento e identificação à sua audiência.

- a) Semeadura: agentes epistêmicos perniciosos considerados confiáveis em determinado grupo intencionalmente enviam, compartilham ou apontam estrategicamente notícias enganadoras nas mídias sociais, deixando ocultos seus interesses e objetivos;
- b) Ecoamento: a desinformação é disseminada nas redes por meio de controvérsias preexistentes, contrapondo retoricamente os valores e crenças identitárias a narrativas, crenças adversárias e a opositores criados por esses agentes perniciosos.

Assim, a confiabilidade e a coesão entre os membros do grupo aumentam (efeito semelhante à corroboração retroalimentativa), e a confiança de quem não é membro e em autoridades epistêmicas sobre os temas e conteúdos sob os quais os grupos são constituídos e têm crenças contrárias reduz-se (Ruiz; Nilsson, 2022), como a investigação que trata de terraplanistas, seus algozes epistêmicos são geólogos e físicos, por exemplo – aumentando a polarização entre quem participa desse grupo e quem não participa, dificultando aos agentes epistêmicos desse ambiente o acesso e a circulação de informações relevantes para a quebra desse círculo vicioso, tornando esses ambientes epistemicamente vulneráveis, como é destacado em Sullivan *et al.* (2020).

Desta forma, vemos aqui pontos em que o fenômeno da câmara de eco é extremamente prejudicial à nossa vida social cada vez mais dependente de interações mediadas por sistemas de informação em que a circulação de informação como um todo está sendo intencionalmente corrompida por notícias falsas e danosas às nossas instituições políticas e econômicas. As bolhas epistêmicas, como relatadas na literatura, têm graus mais atenuados de nocividade, devido a sua natureza menos dependente da manutenção de uma rede de crenças por agentes epistêmicos com interesses escusos e a ocorrência de exposição dos agentes à diversidade de opiniões e informações. Contudo, nada impede – na verdade as mídias até têm vieses algorítmicos para que isso ocorra – que uma bolha epistêmica torne-se uma câmara de eco, portanto mais prejudicial. Mesmo considerando alguns filósofos que não considerem tais fenômenos epistemicamente problemáticos *per se* – como Lackey (2021)¹³ – mas como problemática a confiabilidade do conteúdo informativo que é circulado nesses ambientes, esse aspecto de desconfiança e falta de fidedignidade veritativa da informação deve ser observado seriamente.

Portanto, devemos buscar sanar ao menos essa falta de confiabilidade informativa contida nas redes manifestada por ambos os fenômenos. No entanto, como fazer com que

¹³ Para um ponto de vista contrário, ou seja, argumentando em defesa de um problema epistemicamente intrínseco das bolhas epistêmicas e câmaras de eco (relacionado à funcionalidade dos ambientes virtuais em que tais fenômenos se manifestam e o compromisso social em dever manter-se neles ou não), ver Sheeks (2023).

esses fenômenos sociais sejam mitigados? Há certas soluções sendo desenvolvidas e implementadas. Falaremos sobre isso nas seções abaixo, dando posterior enfoque a certas características que mantêm e fazem desses ambientes sociais virtuais confiáveis para os agentes epistêmicos inclusos nele. Mas antes, vamos retomar os pontos abordados nas seções anteriores e mostrar soluções propostas até aqui.

4 FAZENDO BALANÇO, SOLUÇÕES E MÉTODOS PROPOSTOS

Até aqui, abordamos a bibliografia corrente no tema que relaciona a temática da pós-verdade a aspectos epistêmicos das mídias sociais. Porém, precisamos organizar quais as observações foram feitas até aqui para entendermos as propostas existentes no cânone e do que precisamos para dar um passo adiante na prevenção e mitigação do atual espalhamento de má informação ou desinformação. Destacamos ao longo do artigo os principais temas e por meio deles investigaremos soluções que sejam efetivas dentro da literatura, procurando ir um pouco além. Agora, faremos o que é chamado de fazer um balanço (*taking stock*) do que foi abordado. Em resumo, auxiliam na propagação de desinformação e formação de crenças falsas ou perniciosas:

- i. Problemas estruturais de instituições políticas, econômicas e das mídias que embasam e promovem a formação de crenças alinhadas a determinados interesses, redes com viés algorítmico e alavancamento de notícias de baixa credibilidade por *bots* sociais;
- ii. Os já destacados vieses cognitivos, como lacunas e falhas de raciocínio e argumentação críticas, bem como o recurso a heurísticas e insensibilidade ao conteúdo informacional;
- iii. Elevada atribuição de confiabilidade, credibilidade e até dependência por parte do agente epistêmico relativo ao grupo social a que ele pertence em bolhas epistêmicas e câmaras de eco;
- iv. Realçamento de aspectos afetivos, suscitando sentimentos de pertencimento e identidade e reações intensas dos agentes epistêmicos, culminando em insatisfação e reclamações.

A partir dessas constatações, há na literatura uma gama de propostas. O artigo Pennycook e Rand (2021) fala sobre abordagens correntes para evitar a propagação de má-informação, como de identificação de conteúdo problemático utilizando aprendizado de máquina, processamento de linguagem natural e análise de rede, profissionais realizando

checagem de fatos para anexarem avisos em notícias falsas¹⁴ ou enfatizar editores de artigos de notícias. Essas três formas têm certos problemas, sendo a segunda mais efetiva e destacada em conjunto com novas abordagens, como de conscientização do público por meio de jogos tutoriais contra má-informação e avaliações de precisão de notícias neutras para aumentar as habilidades e conhecimentos desses possíveis agentes epistêmicos, bem como a utilização por parte das plataformas de aprendizado de máquina para capturar uma grande quantidade de avaliações de confiabilidade do público leigo geral (sem precisar recorrer a profissionais), com base na teoria da “sabedoria das multidões”, como mostrado em Allen, J. *et al.* (2020). Essas abordagens serão discutidas criticamente na próxima seção.

Já no problema exposto por Franke (2022), o viés algorítmico político pode ser prejudicial por haver certa conformação por parte dos usuários acerca de como é classificado¹⁵ politicamente e, portanto, aumentam-se o entrincheiramento e a polarização das mídias sociais (Peters, 2022), alegando que essa tese pode ter o resultado contrário ao esperado devido ao fenômeno chamado de efeito laço (*looping effect*)¹⁶ em classificações dentro de tipos das ciências sociais, fazendo com que os indivíduos que se enquadram na classificação dada pelo algoritmo acabam modificando seus comportamentos, crenças e ações devido às consequências de tal ato de classificação e, por isso, os agentes podem na verdade se despolarizar ou despolitizar de certas bolhas que possam vir a ter algoritmos enviesados, buscando imparcialidade e desvinculação de rótulos políticos. Um estudo empírico mais amplo seria interessante para corroborar tal tese.

Algumas análises propositivas já envolvendo a mitigação ou dissolução de bolhas e câmaras de eco têm base no desenvolvimento de pesquisas dentro da epistemologia social. Por exemplo, Nguyen (2018) propõe um reinício social epistêmico (*social epistemic reboot*) no sentido de o agente fazer o exercício intelectual de considerar outras crenças sem qualquer privilégio epistêmico por indivíduos e grupos, levando em conta de forma equânime todo testemunho e suspendendo as crenças que a princípio tinham credibilidade garantida. Em

¹⁴ Esse mesmo dispositivo está em Achimescu e Chachev (2021). Uma abordagem preventiva semelhante de correção factual é destacada em Vosougui, Roy e Aral (2018).

¹⁵ Ele se baseia na maneira distintiva que o filósofo canadense Ian Hacking (1999) propõe para classificar e nomear entidades nas ciências naturais e sociais, denominando-os de tipos naturais e tipos humanos, respectivamente. Mais para frente em outra obra, Hacking acrescenta à noção de tipo humano a de tipo interativo, na qual ocorre o efeito laço.

¹⁶ Podemos ter tanto tipos humanos que apresentem as mesmas características de tipos naturais no surgimento e aplicação do conceito referente a ele, quanto outro tipo que gera modificações nos próprios particulares em que a conceitualização altera o que é tipificado, chamado de tipo interativo, e a seu fenômeno dinâmico e iterativo dá-se o nome de efeito laço. Hacking (2009, p. 63) especifica um tipo interativo quando certas categorias de pessoas “[...] passam a existir no mesmo momento que tipos de pessoas passam a existir de modo a se encaixarem nessas categorias, e há uma interação de mão dupla entre esses processos”.

Boyd (2018), destaca-se a relevância de identificar um grupo epistemicamente pernicioso, e essa identificação está associada com as consequências do problema do agrupamento, em que temos um grupo visto como tendo alta credibilidade pelos seus integrantes – estes inclusive tomando tal grupo como fonte de informação – sendo seus agentes epistêmicos, individualmente, de baixa credibilidade. Há também em epistemologia computacional a implementação de modelos para avaliar conceitos epistêmicos que visem o entendimento desses ambientes de propagação de notícias falsas, como a associação deles com as noções de segurança e vulnerabilidade epistêmica (Sullivan *et al.*, 2020) e o papel da noção de imparcialidade como atitude doxástica de baixo entrincheiramento, representando uma posição epistêmica de máxima persuasão (Baumgaertner, 2014).

Ao final, temos em Ruiz e Nilsson (2022) a identificação de que a circulação de desinformação está ligada a epistemologias muito frágeis e contraditórias entre si (como a de apelo às escrituras sagradas ou um tipo de empirismo ingênuo baseado na experiência direta por terraplanistas, por exemplo). Logo, dadas as fases de semeadura e ecoamento das câmaras de eco, são necessárias medidas distintas nas fases. Na de semeadura, eles incentivam muitas das medidas já destacadas acima no nosso artigo, contudo no ecoamento a estratégia muda, pois o caráter argumentativo torna-se importante na criação de controvérsias e espalhamento da informação em seu ambiente epistêmico. Portanto, é preciso duas estratégias:

- A) Criação de políticas públicas que atinjam a visão de mundo da audiência desejada (nesse caso, agentes epistêmicos em câmaras de eco) com a utilização de instrumentos dentro das próprias epistemologias estabelecidas nas câmaras para realização de experimentos e argumentos contrários às crenças formadas nesses ambientes, e;
- B) Tais experimentos e argumentos devam ser divulgados por aliados, agentes epistêmicos que possam compartilhar certas crenças comuns entre os integrantes da câmara, mais propícios a ter sua confiança.

A citação exemplar abaixo deixa bem claro o objetivo:

[Aqueles que] fazem políticas públicas não podem confiar exclusivamente em declarações factuais proferidas por figuras de autoridade, porque lhes falta *ethos* ou credibilidade dentro da câmara de eco. Em vez disso, a mensagem deve vir de aliados das controvérsias motivadas por identidade. Por exemplo, os defensores da Terra plana argumentam que os bons cristãos devem acreditar na Terra plana; caso contrário, eles são hereges. Para contrariar esse argumento, apenas um cristão genuíno pode dizer que alguém pode permanecer um cristão devoto sem ser um defensor da terra plana (Ruiz; Nilsson, 2022, p. 32, tradução nossa).

Dessa maneira, expomos nesta etapa o apanhado geral de abordagens já existentes na literatura. Observamos que dos quatro itens destacados, (i) trata-se de problemas próprios da constituição e permissões de instituições tradicionais e plataformas de mídias sociais. Mas (ii), (iii) e (iv) são problemas fundamentais que se encontram em uma área de investigação: técnicas de argumentação e retórica, ou seja, pragmática lógica – sendo que em poucos artigos isso é destacado implícita ou explicitamente.

Na próxima seção, elaboraremos nossa crítica em que praticamente todas essas tentativas, mesmo que de início efetivas, sozinhas encontram certas limitações ao serem implementadas, e a partir do reconhecimento da análise de Ruiz e Nilsson (2022), proporemos duas medidas (uma de longo prazo e outra de médio e curto prazo) que, ao nosso ver, poderão preencher as lacunas deixadas.

6 CRÍTICAS E PROPOSTAS: SOBRE EDUCAÇÃO INFORMACIONAL E GRAUS DE PERSUASÃO

Chegamos até aqui buscando traçar o quadro geral dos problemas e soluções da seção anterior. Agora vamos nos deparar com certas dificuldades na implementação dos métodos e ideias acima. Ora, Pennycook e Rand (2021) destacam algumas abordagens que necessitam de um público que ao menos seja minimamente imparcial epistemicamente, tendo receptividade ao aprendizado e à correção por alguma autoridade epistêmica e que não ponha em suspeita nem as intenções dessa autoridade nem a plataforma em que estão sendo implementadas tais medidas – como no caso da checagem de fatos com aviso de conteúdo em conjunto com conscientização e treinamento para identificação de notícias falsas.

Dessa forma, até a análise de confiabilidade com base em sabedoria das multidões envolve tal problema: estudos que consigam identificar ambientes de que possamos extrair informações características de sabedoria das multidões – como diversidade de opinião e descentralidade de decisões – e de diferenciá-los de seu contraposto epistêmico – a loucura das massas, como destacado em Sullivan *et al.* (2020, p. 3) – também devem lidar com o fato de que tais valores podem ser considerados politicamente enviesados por pessoas que integram ambientes perniciosos – como as bolhas e câmaras já citadas – sendo justamente as os agentes que se encontram nesses ambientes parte relevante do público-alvo dessas medidas. Podemos chamar isso de um tipo de viés cético ou viés de valoração carregada (*value-laden bias*): agentes epistêmicos que detêm um sistema de crenças baseado na negação

e distanciamento de certos valores que embasam a implementação de certos métodos ou políticas dificilmente irão aceitar esses métodos ou políticas como legítimos.

O reinício social epistêmico proposto em Nguyen (2018), inspirado na dúvida hiperbólica cartesiana, parece uma visão meio dúbia e um tanto subjetiva de como escapar de uma câmara de eco. Depender de uma capacidade reflexiva que está intrinsecamente associada ao cultivo de virtudes intelectuais como a autonomia, sendo que o indivíduo se encontra cooptado por um ambiente virtual em que virtudes como essas são solapadas com o forte estabelecimento de confiança mútua epistemicamente codependente dentro na câmara não faz sentido e soa contraditório, aparentando mais uma solução motivada por esperança que uma proposta realmente eficaz. Mesmo havendo exemplos radicais desse reinício, basear-se em exceções é, novamente, característico de atitudes esperançosas que não deveriam ser um caminho por onde nortear-se.

Em suma, mesmo tais métodos e ideias sendo efetivos, temos que contornar possíveis problemas envolvidos em suas implementações, buscando soluções de longo prazo. Quando falamos em um período de tempo razoável para estabelecer certo conhecimento técnico ou comportamento social, um planejamento público por meio de campanhas seria desejável (tal como as campanhas pelo uso de cinto de segurança, por exemplo). Porém, nossa sociedade sofreu grandes transformações com as Tecnologias de Comunicação e Informação (ICT, em inglês) e ainda está se organizando para entender como, no nível educacional básico, as ICT devem integrar instituições governamentais e educacionais. A maioria das abordagens se refere à aplicação dessas tecnologias e sobre a informatização de serviços institucionais, porém pouco material é produzido sobre disciplinas de ICT no currículo básico de escolas, mesmo a própria UNESCO tendo um programa de ensino já em 2002 (Anderson; Van Weert; Duchâteau, 2002). ICT não apenas como instrumento para um aprendizado personalizado (Assar, 2015), mas como de fato conteúdo disciplinar. Porém, as ICT mudaram muito de 2002 até os dias de hoje.

Os dispositivos eletrônicos com plataformas de *software* com acesso a Internet são os primeiros artefatos técnicos que passaram a mediar profundamente – com sérias implicações sociais e até filosóficas (Verbeek, 2015) – nossa experiência no cotidiano de forma contínua e interativa, nos obrigando a aprender seu manuseio. Podemos utilizar esses dispositivos sem saber sua constituição, tal como óculos de grau e veículos automotores. No caso dos veículos, podemos nos deslocar com eles sem saber dirigir; há especialistas que os conduzem, e não precisamos aprender manuseio nem regras específicas e ética no trânsito (mesmo de que

regras voltadas à interação de carros e pedestres precisemos). Tornou-se praticamente impossível utilizar tais dispositivos sem adentrar em um ambiente epistêmico que nos coloque em uma situação como se estivessem constituindo paulatinamente parte de nossa própria identidade (Floridi, 2014, cap. 3). Faz-se necessário mais do que campanhas de conscientização: devemos elaborar políticas públicas de institucionalização educacional do aprendizado e conscientização acerca dos danos de conteúdos desinformativos na sociedade, sua identificação e a de ambientes epistêmicos virtuais nos quais sua propagação é mais efusiva¹⁷. Disciplinas de ICT em instituições educacionais seria o ambiente epistêmico ideal, tanto pela minimização de resistência à adesão das dinâmicas e do conteúdo educativo quanto pela melhor aceitação de legitimidade das autoridades epistêmicas envolvidas. No Brasil essa abordagem encontra-se de forma vaga ou nas entrelinhas de diretrizes e habilidades em certas áreas no BNCC, sendo curricularmente pouco relevante e subdeterminada¹⁸.

Claramente ainda precisamos lidar com quem se encontra nas câmaras de eco e filtros de bolha, que já se encontram imersos em ambientes epistemicamente perniciosos, propagando desinformação e formando crenças reforçando-as pelos vieses cognitivos, processos de agrupamento e corroboração retroalimentada – e mesmo quando convencidos de alguma falsidade, mantendo ecos de crença. Precisamos, nesse caso, procurar fazer uso desses fenômenos e efeitos para alcançar a visão de mundo desses indivíduos. E temos em Ruiz e Nilsson (2022) um ótimo *insight* baseado nas epistemologias contraditórias destes grupos para criar iniciativas a persuadi-los.

Dessa forma, podemos avançar as ideias de (A) e (B) – esta última direcionada também por ideias do artigo em Retórica – pensando nos graus de separação que temos de outros agentes epistêmicos: compartilhamos várias crenças em comum com pessoas que se encontram parcialmente nos mesmos grupos que nós, e essas outras pessoas encontram-se em

¹⁷ Esse tipo de disciplina poderia ter benefícios até para o gerenciamento pessoal de ambientes com excessos de informação com os quais nos deparamos e não conseguimos lidar, como em mangueiras de falsidades (*firehose of falsehoods*) – tal fenômeno é definido como enxurrada epistêmica (*epistemic flooding*) (Anderau, 2023, p. 108).

¹⁸ No BNCC, já nas competências gerais da Educação Básica, a de número cinco reforça “compreender, utilizar e criar tecnologias digitais de informação e comunicação de forma crítica, significativa, reflexiva e ética nas diversas práticas sociais (incluindo as escolares)” (Brasil. MEC, 2018, p. 9) e elaborando sobre as ICT, chamadas de Tecnologias Digitais Informação e Comunicação (TDIC), destacando “[...] as aprendizagens relativas às formas de processar, transmitir e distribuir a informação de maneira segura e confiável em diferentes artefatos digitais – tanto físicos (computadores, celulares, *tablets* etc.) como virtuais (internet, redes sociais e nuvens de dados, entre outros)” (Brasil. MEC, 2018, p. 474), direcionando esses objetivos e competências da área de Linguagens, especificamente na competência específica seis e sete no Ensino Fundamental e Médio, respectivamente, relacionando-se este último ao campo de atuação social jornalístico-midiático. Porém, mesmo com alguma discussão sendo promovida sobre ética e confiabilidade da informação, não há em nenhum momento referências a mitigar propagação de desinformação.

uma circunstância social semelhante, e assim por diante. Depois de alguns poucos graus de separação, talvez teremos uma pessoa que não compartilhe crença alguma com o primeiro. Se fôssemos colocá-los para debater acerca de um tema em que haja discordância, quanto menos crenças as pessoas em um debate compartilham, mais difíceis serão suas concessões na discussão; a persuasão não será muito efetiva e, muito provavelmente, a conversa não será epistemicamente proveitosa. Porém, se tomarmos alguém com grau de separação no ponto de vista epistêmico menor, a persuasão pode ser maior e a possibilidade de convencimento mais provável. Portanto, muitas vezes uma abordagem argumentativa indireta – por meio de agentes epistêmicos com um sistema de crenças similar o suficiente para haver um diálogo construtivo, porém distinto o suficiente para que tragam ao diálogo o componente doxástico revisivo – seja um método mais efetivo de visar à revisão de opiniões de agentes em grupos epistemicamente perniciosos como os de Boyd (2018): temos aqui um princípio de revisão de crenças com base em uma teoria retórica de níveis ou graus de persuasão e possibilidade de convencimento.

Aqui convém esclarecer que estão sendo tomados persuasão e convencimento como conceitos distintos. Tal como é explicado em Cattani (2020), persuadir é uma noção mais centrada no processo de argumentação retórica e a uma capacidade do falante de afetar o comportamento e ações do ouvinte, enquanto o convencimento seria um estado mental – do ouvinte ou plateia – associado à racionalidade e à evidência que se apresenta no argumento. Assim, o ato de dissuadir um ouvinte (o antônimo de persuadir) faz com que ele mude seu comportamento e certas crenças, mas não significa que ele se convenceu, mesmo que em um debate a persuasão seja necessária ao convencimento. Perelman e Olbrechts-Tyteca (1996, p. 30) destacam a relevância de cada um, dependendo da preocupação em foco:

Para quem se preocupa com o resultado, persuadir é mais importante que convencer, pois a convicção não passa da primeira fase que leva a ação. Para Rousseau, de nada adianta convencer uma criança “se não se sabe persuadi-la”. Em contrapartida, para quem está preocupado com o caráter racional da adesão, convencer é mais importante do que persuadir. Aliás, ora esta característica racional da convicção depende dos meios utilizados, ora das faculdades às quais o orador se dirige.

Claramente, estamos mais preocupados com aspectos persuasivos aqui, pois a nossa intenção principal é propor medidas para mitigar a propagação de desinformação – se um ato dissuasivo de qualquer falante fizer com que seu ouvinte não tome iniciativas de compartilhar ou divulgar notícias falsas, já temos um avanço em tornar nossos ambientes epistêmicos mais saudáveis. Porém, se obtivermos o convencimento desses ouvintes, melhor seria. Como em

parte destacamos acima, a iniciativa é pensar em graus de persuasão por meio das inter-relações dos sistemas de crenças (tais crenças implícitas ou explícitas) e dos graus de separação entre agentes em um ambiente epistêmico. Quando falamos em sistema de crenças, estamos nos referindo desde a afirmações sobre fatos até conjuntos de valores intersubjetivos – como no exemplo em Ruiz e Nilsson (2022) de ser um “cristão genuíno”. Destarte, queremos com isso abarcar retoricamente aspectos do *ethos* (ambiente epistêmico e relações entre agentes – confiabilidade, por exemplo), *pathos* (conteúdos afetivos envolvidos nos valores, sistemas de crenças inter-relacionados e relações entre agentes) e *logos* (conteúdo informacional nos sistemas de crenças inter-relacionados) e afirmar que, significativamente, dependendo de tais fatores, *certos agentes epistêmicos possam ter maior grau de persuasão para com determinado ouvinte/plateia que outros*, i.e., podemos por meio de alguns elementos retóricos estabelecer o grau persuasivo entre agentes epistêmicos. No caso do convencimento essas variáveis também seriam consideradas para estipularmos o potencial de convencimento, porém haveria uma maior relevância do *logos* em sua quantificação – seja ela qual for.

Além de utilizarmos essa base teórica para estipularmos o agente epistêmico mais adequado para persuadir outro, temos também outro aspecto a ser considerado: a estrutura do argumento e seu conteúdo informativo. Vimos nas seções acima que, dependendo da mídia e público a ser alcançado, há maneiras distintas com base em epistemologias distintas que elevam a importância e a estima dos agentes epistêmicos que estamos considerando com relação à informação propagada. Alegamos que, por vezes, para defendermos determinada tese argumentativamente, ela precisa ser transmitida de maneiras persuasivamente diretas ou indiretas. Aqui recorremos a Pinto (2001) e McKeon (2020), em que temos no primeiro a noção de argumentação como um convite à inferência, culminando na elaboração pelo último de que certas argumentações induzem o agente de forma inferencialmente indireta. Convite a uma inferência que induz a crença, tal que “[...] possa persuadir indiretamente um destinatário [ouvinte], servindo-o como um guia à performance de uma inferência, mesmo que seu conteúdo não seja expresso pelo argumento, ou a conclusão não seja esboçada da forma pretendida pelo argumentador” (McKeon, 2020, p. 349, tradução nossa).

A esse adendo no final da citação podemos contrastar com McKeon que aquilo que almejamos na persuasão indireta aqui pode ser de caráter intencional, exemplo: quando eu estou debatendo sobre não comer carne animal com base no argumento de senciência em animais vertebrados com uma pessoa que tenha razoável resistência em ser persuadida no

tema, posso ao invés de falar acerca da complexidade neurológica de vertebrados proporcionar-lhes desenvolvimento cognitivo, falar sobre experiências em comum com a pessoa relativo ao comportamento intencional de animais domésticos, como eles têm sonhos, parecem saber quando fazem algo que não deveriam etc. Depois, trazer informações sobre a semelhança entre mamíferos superiores domésticos e domesticados utilizados como alimento. Há grandes chances de essa forma de persuasão indireta envolvendo suas experiências diretas com animais domésticos com os quais já existe apego afetivo a familiaridade (*ethos e pathos*) trazendo a semelhança entre mamíferos superiores (*logos*) elevar o grau de persuasão à afirmação principal: não comer carne animal. E nesse exemplo estamos pensando em grau de separação zero, ou seja, estamos falando diretamente com a pessoa. Porém, podemos pensar casos em que o grau de persuasão é elevado quando consideramos afetar o comportamento de outros agentes que podem persuadir o meu agente epistêmico-alvo de forma mais efetiva. Se quero chamar a atenção de um agente P acerca de um comportamento reprovável seu, porém meu grau de persuasão com relação a ele é baixo, mas tenho grau considerável com um amigo Q que por vezes tem o maior grau de persuasão com relação a P, posso argumentar com Q incentivando ele a argumentar acerca do comportamento reprovável de P com o próprio, e temos neste caso grau de separação um ao ouvinte/agente epistêmico-alvo.

Outros dois exemplos mais completos e associados à aplicação dessa teoria às mídias sociais: vamos supor que muita desinformação esteja sendo difundida nas redes por um site que um agente epistêmico A, pertencente a uma câmara de eco, considere-o como uma fonte confiável e compartilhe vez por outra os conteúdos desse site. A mídia social tem dispositivos de checagem de fatos de todas as notícias falsas divulgadas por esse site até o momento. Vamos pensar em duas situações: uma em que a plataforma queira aumentar seu grau de persuasão com relação a A de que uma determinada notícia F divulgada por este site é falsa e outra circunstância em que se queira minar a confiabilidade de A no site. No primeiro caso, se A tem uma confiabilidade alta no site e pouca na plataforma, há a possibilidade da checagem de fatos ser colocada em xeque por viés de valoração carregada e reforço de crença na desinformação por *backfire effect*. Portanto, pela nossa teoria, podemos difundir a checagem entre outros agentes que tenham bom grau de persuasão com A, levando em conta elementos destacados acima, e que esses agentes considerem tal desinformação F de fato falsa e possam compartilhar tal checagem. O agente epistêmico A, a partir desses compartilhamentos por

parte de agentes persuasivos e confiáveis, tenderá a ser persuadido de que não F. Tal como em Moussaïd *et al.* (2013), agentes funcionarão como atratores da opinião de A¹⁹.

Já no segundo caso, a ideia de desacreditizar o site passaria por reforço sistemático por meio de vários agentes epistêmicos, divulgando-se em suas mídias que várias notícias daquele site são falsas. Assim, podemos selecionar agentes persuasivos com sensibilidade a conteúdos distintos desinformativos do mesmo site e com bom grau persuasivo com relação a A (ou a outros agentes que tenham tal grau persuasivo, apelando ao aspecto de graus de separação) e possam compartilhar as checagens de fatos. Conforme A observa a corroboração das checagens de várias falsidades associadas ao site propagador de desinformação, indiretamente convida-se A a inferir que tal site não é confiável. Essa estratégia serve como estratégia preventiva, pois se A tiver acesso a outras notícias desse site, a desconfiança nele pode levar A a previamente pesquisar sobre a notícia em vez de aceitá-la como verdadeira, levando essas posturas para dentro da câmara de eco. Essas estratégias podem ser implementadas e efetivadas em conjunto com ferramentas conceituais e de análise, aperfeiçoando a noção de máxima persuasão (Baumgaertner, 2014), trabalhando com a concepção de segurança e vulnerabilidade epistêmica (Sullivan *et al.*, 2020) para identificar possíveis ambientes epistemicamente perniciosos em mídias sociais que sejam passíveis da aplicação desse método e solapando fases de ecoamento em câmaras de eco, por exemplo.

Há ainda muito o que ser elaborado na parte teórica e prática nessa abordagem por graus de persuasão, contudo, pode vir a se tornar um método bem frutífero. A seguir, destacamos como isso pode ocorrer e quais linhas de pesquisa seriam promissoras para seu desenvolvimento.

CONCLUSÃO

Ao longo deste artigo, vimos que as concepções de caráter estrutural que vieram a configurar o fenômeno da pós-verdade já assolam a sociedade contemporânea ao menos desde o fim da Primeira Guerra Mundial, com o advento da sociedade de consumo e da propaganda como instrumento de manipulação da opinião e controle de populações, tanto em regimes totalitários quanto democráticos. Há vários fatores que colaboram para a desvalorização da noção de factualidade e da verdade no panorama contemporâneo, tendo como um dos

¹⁹ Ou ao menos influenciarão no comportamento dissuasivo de não compartilhamento da notícia falsa, o que já pode auxiliar no controle de maus hábitos de compartilhamento em mídias sociais (Ceylan; Anderson; Wood, 2023).

principais – indo de acordo com Arendt – a dificuldade política de lidar com tais conceitos, sendo a crença e a opinião mais flexíveis e politicamente manejáveis, e com o tempo ganhando mais proeminência em conjunto com suas cargas afetivas subjacentes. Isso dá abertura à enganação e à propagação de desinformação, e em uma era de informação em que seus critérios de credibilidade se dissociaram de uma autoridade epistêmica e passaram a se basear na confiabilidade dos agentes e seus ambientes, a Web 2,0 em conjunto com as mídias sociais mudaram a configuração epistêmica das ICT.

Há diversas formas de propagação de desinformação nas redes, desde *bots* sociais a criação de ambientes nos quais grupos intencionalmente tornam-se epistemicamente perniciosos, como as câmaras de eco, em que operam agentes mal intencionados retroalimentando a corroboração de crenças baseadas nas propagações de desinformação promovida por tais agentes dentro das mídias sociais, auxiliados por vieses cognitivos característicos de raciocínios falhos, mal informados e epistemicamente pouco virtuosos dos agentes que integram tais plataformas. Vimos as constituições desses grupos e suas consequências, bem como as soluções propostas na literatura para buscarmos mitigar estes fenômenos, em que desenvolvemos críticas e elaborações para aperfeiçoar os métodos e utilizá-los, evitando problemas em suas implementações. Existem certas dificuldades envolvendo as nossas duas contribuições: o foco em políticas educacionais necessita de uma reestruturação do currículo escolar e investimentos. Pensar em maneiras de persuasão de estratégia preventiva voltadas à aceitação desta proposta é um trabalho em aberto.

Assim como nossa segunda proposta, que necessita de formas de análise de dados e sua quantificação para definir a melhor forma de estipular os graus de persuasão e sua implementabilidade, em Zachlod *et al.*, (2022) há estudos sobre análise e monitoramento de dados em mídias sociais e suas aplicações, e no artigo de Shu, Bernard e Liu (2018) já há alguns avanços relevantes em detecção de notícias falsas, propagação de credibilidade e até identificação de agentes persuasivos; como aplicado no Twitter (Schaefer; Stede, 2021), temos que a recente área de mineração de argumentos pode, através de *datasets*, técnicas de redes neurais e aprendizagem profunda (*Deep Learning*), extrair e avaliar argumentos nas mídias sociais com intuito de quantificar a variabilidade de graus de persuasão entre diferentes apresentações de argumentos; já dentro da Epistemologia Social e Computacional, temos a abordagem de modelos computacionais – tal como o modelo de crença tricotômico (Baumgaertner, 2014) – rodando simulações para entender o desempenho de nosso método de graus de persuasão em ambientes epistêmicos distintos. Tais investigações futuras, caso sejam

prolíficas, poderão nos esclarecer e auxiliar a impedir essa cooptação de pessoas dentro de ambientes epistêmicos cognitivamente nocivos e de propagação excessiva de desinformação nas mídias sociais.

REFERÊNCIAS

- ACHIMESCU, V.; CHACHEV, P. D. Raising the flag: monitoring using perceiving disinformation on Reddit. **Information** 12(1): 4, 2021.
- ALLEN, J. N. L.; ARECHAR, A. A.; PENNYCOOK, G.; RAND, D.G. Scaling up fact-checking using the wisdom of crowds. **PsyArXiv Published online**, 2020.
- ANDERAU, Glenn. Fake News and epistemic flooding. **Synthese** 202:106, 2023.
- ANDERSON, J.; VAN WEERT, T.; DUCHÂTEAU, C. (ed.). **Information and communication technology in education: a curriculum for schools and programme of teacher development**. Paris: UNESCO, 2002.
- ARENDT, Hannah. **Entre o passado e o futuro**. Tradução de Mauro Barbosa. São Paulo: Editora Perspectiva, 2011.
- ARENDT, Hannah. Lying in politics: reflects on the Pentagon Papers. *In*: ARENDT, Hannah. **Crises of republic**. Nova York, Harcourt Brace & Company, 1972. p. 1-48.
- ASSAR, Saïd. Information and communications technology and education. *In*: Wright, James D. **International Encyclopedia of the Social & Behavioral Sciences** (second edition), Elsevier, 2015. p. 66-71.
- BAUMGAERTNER, B. Yes, no, maybe so: a veritistic approach to echo chambers using a trichotomous belief model. **Synthese** 191, p. 2549-2569, 2014.
- BENNETT, W. L.; LIVINGSTON, S. A brief story of disinformation age: information wars and the decline of institutional authority. *In*: BENNETT, W. L.; LIVINGSTON, S. (ed.) **The disinformation age: politics, technology and disruptive communication in the United States**. Cambridge: University Press, 2021. p. 3-42
- BOYD, Kennedy. Epistemically Pernicious Groups and the Groupstrapping Problem. **Social Epistemology**, 33 (1), p. 61-73, 2018.
- BRASIL. Ministério da Educação. **Base Nacional Comum Curricular**. Brasília: MEC, 2018.
- CATTANI, A. **Persuading and convincing**. Ontario Society for the Study of Argumentation Conference Archive 11, 2020.
- CEYLAN, G.; ANDERSON I. A.; WOOD, W. **Sharing of misinformation is habitual, not just lazy or biased**. *Proc Natl Acad Sci USA* 120(4): e2216614120, 2023.

CINELLI, M.; MORALES, Gianmarco de Francisci; GALEAZZI, A.; QUATTROCIOCCHI, W.; STARNINI, M. **The echo chamber effect on social media**. Proceedings of the National Academy of Sciences 118(9): e2023301118, 2021.

DUTILH MORAES, C.; DE RIDDER, J. Is Fake News old News? *In*: BERNECKER, S., FLOWERREE, A. K. and GRUNDMANN, T. (ed.). **The epistemology of Fake News**. New York, NY: Oxford University Press, 2021.

EMPOLI, Giuliano da. **Engenheiros do caos**. Tradução de Arnaldo Bloch. São Paulo: Editora Vestígio, 2019.

ENDERS, A. M.; USCINSKI, J. E.; SEELIG, M. I.; KLOFSTAD, C. A.; WUCHTY, S.; FUNCHION, J. R.; STOLER, J. The relationship between social media use and beliefs in conspiracy theories and misinformation. **Political Behavior**, 45: p. 781-804, 2021.

FEIGOLD, S. **The four key ways desinformation is spread online**. World Economic Forum, 2022.

FERRARA, E.; VAROL, O.; DAVIS, C.; MENCZER, F.; FLAMMINI, A. **The rise of social bots**. Communications of the ACM, vol. 59, no. 7, p. 96-104, 2016.

FLORIDI, Luciano. **The fourth revolution: how the infosphere is reshaping human reality**. Nova York: Oxford University Press, 2014.

FLORIDI, Luciano. Prominence-interpretation theory: Explaining how people assess credibility online. *In*: COCKTON, G.; KORHONEN, P. (ed.). **Proceedings of the SIGCHI Conference on human factors in computing systems**. Flórida: Fort Lauderdale, 2003.

FOGG, B. J. **Stanford guidelines for web credibility**. A research summary from the Stanford Persuasive Technology Lab. Stanford University, 2002.

FRANKE, U. Algorithmic political bias – an entrenchment concern. **Philosophy and Technology**, 35 (3): p. 1-6, 2022.

GILES, Jim. Internet encyclopaedias go head to head. **Nature**, v. 438, n. 7070, p. 900-901, 2005.

HACKING, Ian. **The social construction of what?** Cambridge, Mass: Harvard University Press, 1999.

HACKING, Ian. **Ontologia histórica**. Rio Grande do Sul: Editora Unisinos, 2009.

HIMELEIN-WACHOWIAK, M; GIORGI S.; DEVOTO, A; RAHMAN, M.; UNGAR L.; SCHWARTZ, H.A.; EPSTEIN, D.H.; LEGGIO, L.; CURTIS, B. Bots and Misinformation Spread on Social Media: Implications for COVID-19. **J Med Internet Res** 23(5): e26933, 2020.

JAMIESON, Kathleen H.; CAPPELLA, Joseph N. **Echo chamber: rush limbaugh and the conservative media establishment**. New York: Oxford University Press, 2008.

KEYES, Ralph. **A era da pós-verdade**: desonestidade e enganação na vida contemporânea. Tradução de Fábio Creder. Petrópolis, RJ: Editora Vozes, 2018.

LACKEY, Jennifer. Echo Chambers, Fake News and social epistemology. *In*: BERNECKER, S., FLOWERREE, A.K. and GRUNDMANN, T. (eds.). **The Epistemology of Fake News**, p. 206-227. New York, NY: Oxford University Press, 2021.

LANKES, R. D. Credibility on the internet: shifting from authority to reliability. **Journal of Documentation**, v. 64, Issue: p. 667-686, 2008.

MCKEON, Matthew W. Argument, inference, and persuasion. **Argumentation** 35 (2): p. 339-356, 2020.

MOUSSAÏD, M.; KÄMMER, J. E.; ANALYTIS, P. P.; NETH, H. Social influence and the collective dynamics of opinion formation. **PLoS One**, 8(11): e78433, 2013.

NGUYEN, C. Thi. Echo chambers and epistemic bubbles. **Episteme** 17(2): p. 41-61, 2020.

PANSANELLA, V.; SÎRBU, A.; KERTÉSZ, J.; ROSSETTI, G. Mass media impact on opinion evolution in biased digital environments: a bounded confidence model. **Sci Rep**. 13(1): 14600, 2023.

PARISER, Eli. **The filter bubble**: what the internet is hiding from you. New York: Penguin Books, 2011.

PENNYCOOK, G; RAND, David G. The Psychology of Fake News. **Trends in Cognitive Sciences**, v. 25, Issue 5, p. 388-402, 2021.

PERALTA, A. F.; NERI, M.; KERTÉSZ, J.; IÑIGUEZ, G. Effect of algorithmic bias and network structure on coexistence, consensus, and polarization of opinions. **Phys Rev E**. 104(4-1): 044312, 2021.

PERELMAN, C.; OLBRECHTS-TYTECA, L. **Tratado de argumentação**: a nova retórica. Tradução Maria Ermantina Ferreira. São Paulo: Martins Fontes, 1996.

PETERS, U. Algorithmic political bias can reduce political polarization. **Philosophy & Technology** 35, 81, 2022.

PILLAI, Rainak M.; FAZIO, Lisa K. The effects of repeating false and misleading information on belief. **WIREs Cogn Sci**; e1573, 2021.

PINTO, Robert C. The relation of argument to inference. *In*: PINTO, Robert C. (ed.). **Argument, inference, and dialectic**. Dordrecht, Kluwer Academic Pub, p. 32-45, 2001.

QUINN, Aaron. Fake News, false beliefs, and the need for truth in journalism. **International Journal of Applied Philosophy** 31 (1): p. 21-29, 2017.

RUIZ, Carlos D.; NILSSON, Tomas. Disinformation and echo chambers: how disinformation circulates on social media through identity-driven controversies. **Journal of Public Policy & Marketing** 42:1; p. 18-35, 2022.

SAMUEL, Lawrence R. **Freud on Madison Avenue**: motivation research and rubliminal rdvertising in America - University of Pennsylvania Press, 2011.

SANTAELLA, Lúcia. **A pós-verdade é verdadeira ou falsa?** Barueri, SP: Estação das Letras e Cores, 2018.

SCHAEFER, R.; STEDE, M. Argument mining on Twitter: a survey. **Information Technology**, v. 63, no. 1, p. 45-58, 2021.

SHAO, C.; CIAMPAGLIA, G. L.; VAROL, O.; YANG, K. C.; FLAMMINI, A.; MENCZER, F. The spread of low-credibility content by social bots. **Nat Commun** 9(1): 4787, 2018.

SHAREVSKI, F.; ALSAADI, R; JACHIM, P.; PIERONI, E. Misinformation warnings: Twitter's soft moderation effects on COVID-19 vaccine belief echoes. **Computers & Security**. v. 114, 102577, 2022.

SHEEKS, Meredith. The myth of the good epistemic bubble. **Episteme** 20 (3): p. 685-700, 2023.

SHU, K.; BERNARD, H. R.; LIU, H. Studying Fake News via Network Analysis: Detection and Mitigation. In: AGARWAL, N.; DOKOOHAKI, N.; TOKDEMIR, S. (ed.). **Emerging research challenges and opportunities in computational social network analysis and mining**, Springer, p. 43-65, 2018.

SULLIVAN, E.; SONDAG, M.; RUTTER, I.; MEULEMANS, W.; CUNNINGHAM, S.; SPECKMANN, B.; ALFANO, M. Vulnerability in Social Epistemic Networks. **International Journal of Philosophical Studies**, p. 1-23, 2020.

SUNDIN, O. Janitors of knowledge: constructing knowledge in the everyday life of Wikipedia editors. **Journal of Documentation**, v. 67, n. 5: p. 840-862, 2011.

TESICH, Steve. A government of lies (political ethics). **The Nation**, Nova Iorque, n. 254, p. 12-13, 1992.

THORSON, Emily A. Belief Echoes: The persistent effects of corrected misinformation. **Political Communication**, 33(3): p. 460-480, 2015.

VERBEEK, Peter-Paul. Beyond interaction: a short introduction to interaction theory. **Interactions** XXII (3): p. 26-31, 2015.

VOSOUGUI, S; ROY, D.; ARAL, S. The spread of true and false news online. **Science** 359, issue 6380: p. 1146-1151, 2018.

WEEDON J.; NULAND, W.; STAMOS, A. **Information operations and Facebook**. Facebook, 27 April, 2017.

ZACHLOD, C.; SAMUEL, O.; OCHSNER, A.; WERTHMULLER, S. Analytics of social media data: state of characteristics and application. **Journal of Bussiness Research** 144: p. 1064-1076, 2022.