FAKE NEWS: UMA REVISÃO COMPREENSIVA E INTERDISCIPLINAR

FAKE NEWS: A COMPREHENSIVE AND INTERDISCIPLINARY SURVEY

Ederval Pablo Ferreira da Cruz
Universidade Estadual Norte Fluminense, UENF
Campos dos Goytacazes, RJ, Brasil
ederval.cruz@gmail.com

Rodolfo Moura Pereira
Universidade Estadual Norte Fluminense, UENF
Campos dos Goytacazes, RJ, Brasil
rodolfom.ifes@gmail.com

Gilberto Mazoco Jubini
Universidade Estadual Norte Fluminense, UENF
Campos dos Goytacazes, RJ, Brasil
gilmazojubini@gmail.com

Lucas Capita Quarto
Universidade Estadual Norte Fluminense, UENF
Campos dos Goytacazes, RJ, Brasil
lcapitaiv@gmail.com

Carlos Henrique Medeiros de Souza
Universidade Estadual Norte Fluminense, UENF
Campos dos Goytacazes, RJ, Brasil
chmsouza@gmail.com

Resumo. Criar notícias falsas com o objetivo de manipular alguém ou um grupo de pessoas, historicamente não é algo recente. Porém, nos últimos anos, a criação de notícias falsas, ou mundialmente conhecidas como *fake news*, tem se intensificado e sua propagação tem sido cada vez mais acelerada. O nível de complexidade em detectar essas *fake news* tem aumentado significativamente com as *deepfakes* e chamado a atenção da comunidade científica, pois é um tópico de pesquisa relevante, dado os seus impactos na sociedade global. Este trabalho, até o momento, é o primeiro em Português, de revisão de literatura de caráter interdisciplinar que envolve *fake news* e as *deepfakes*, seus conceitos, classificações, formas de propagação e as suas relações em diversas áreas do conhecimento. Os resultados mostram que a solução definitiva está longe de ser alcançada e ainda há muito por se fazer em busca do fim ou da minimização dos impactos causados pelas *fake news* e *deepfakes* na sociedade como um todo.

Palavras chave: fake news; deepfake; interdisciplinar; revisão.

Abstract. Creating fake news for the purpose of manipulating someone or a group of people has historically not been new. However, in recent years, the creation of fake news, or known worldwide as fake news, has intensified and its spread has been increasingly accelerated. In addition, the level of complexity in detecting this fake news has also increased significantly with deepfakes. This has attracted the attention of the scientific community, showing that it is a very relevant research topic, given its impacts on global society. This work consists as a first paper in Portuguese, of an interdisciplinary literature review involving fake news and deepfakes, where concepts, their classifications, ways of propagation and the relationship between fake news and deepfakes in different areas of knowledge are presented. The results show that the definitive solution is far from being reached and there is still much to be done in search of an end or minimizing the impacts caused by fake news and deepfakes on society as a whole.

Keywords: fake news; deepfake; interdisciplinary; survey.

INTRODUÇÃO

Não é de hoje que a difusão de notícias falsas faz parte do cotidiano. Após o avanço da tecnologia no século XIX, que proporcionou os meios necessários para o desenvolvimento e consequente alcance dos meios de comunicação, observaram-se os primeiros relatos de notícias capciosas que foram tomadas como verdade pela população.

Na década de 1920, com o advento dos rádios, registrou-se um dos casos primogênitos, onde uma emissora norte-americana criou uma versão audível do livro "A guerra dos mundos", que causou pânico à população da época que, inadvertidamente, acreditou que sua transmissão se tratava de um evento real de invasão de alienígenas na Terra (Tandoc, Lim, & Ling, 2018). Não obstante, neste caso, o objetivo ter sido de promover uma opção de lazer, sem a intenção da emissora de convencer o público de que aquele momento se tratava de uma invasão verdadeira, o evento ficou historicamente marcado pela demonstração do poder malicioso de uma história falsa.

O termo *fake news* se tornou uma "palavra da moda". Enquanto os meios de comunicação proeminentes, como The New York Times, CNN e Buzzfeed News, usaram o termo para designar informações enganosas espalhadas de forma online, o presidente Donald Trump usou o termo como uma designação negativa desses próprios "meios de comunicação convencionais" ou da "grande mídia" durante as eleições americanas de 2016, segundo Farkas & Schou (2018).

Atualmente, as *fake news* possuem uma exposição significativa nos meios de comunicação, facilitadas principalmente pelo alcance da internet e, principalmente, das redes sociais, levando esse fenômeno a afetar diversas instituições sociais, o que causa enorme prejuízo para a população. Um exemplo recente é observado no Brasil, onde a pandemia da COVID-19 causou a morte de milhares de pessoas. Dentre as ações promovidas pelo Ministério da Saúde, o controle da temperatura corporal, por meio do termômetro infravermelho digital apostada para a testa das pessoas, configura-se como uma dessas ações na tentativa de controle da pandemia. Desconsiderando o fato de que o método é questionado por profissionais da saúde que pontuam que nem todos os infectados pelo referido vírus apresentam o sintoma da febre, tal medida foi alvo da promoção da ideia equivocada de que o laser poderia exercer algum tipo de controle, o que levou muitas pessoas a rejeitarem a medição, deixando os comerciantes em situação delicada perante as recomendações em vigor. A solução encontrada foi a medição da temperatura no punho, método que se mostra mais eficiente tendo resultados de medições mais estáveis, conforme o trabalho de Ge *et al.* (2020).

Diante do exposto, mesmo reconhecendo todas as benesses que a internet trouxe consigo ao século XXI, este estudo foca na necessidade de ampliar as discussões sobre as adversidades vinculadas a este seu efeito colateral, notadamente as *fake news.*

Na literatura é possível encontrar alguns trabalhos de revisão de literatura em português envolvendo *fake news* (Pinto & Zanette, 2020; Souza *et al.*, 2020; Leite, Lourenço & Ferreira, 2018). Porém, até o presente momento, segundo pesquisas realizadas, este é o primeiro trabalho, em português, que aborda *fake news* e *deepfake* aprofundando os temas em diversas áreas sob um viés interdisciplinar. Portanto, nossa contribuição para a literatura, através desta pesquisa, situa-se em promover uma revisão da literatura atual que contribua nos debates acerca do fenômeno das *fake news*, com ênfase em sua conceituação, impactos sociais e visão interdisciplinar.

O restante deste artigo está organizado conforme segue. Na próxima seção, apresenta-se a metodologia utilizada no trabalho. Em seguida, serão investigados alguns trabalhos que conceituam *fake news*, assim como as classificações das mesmas. Após, investiga-se uma nova forma de *fake news*: as *deepfakes*. Depois são discutidas as formas de propagação das *fake news* e das notícias reais. Em seu desfecho, são discutidas as *fake news*, *deepfakes* de forma interdisciplinar, envolvendo suas relações com a computação, a psicologia, entre outras áreas.

METODOLOGIA

Para a realização da pesquisa bibliográfica, foi abordada uma pesquisa exploratória utilizando o mecanismo de busca de artigos científicos do Google: o Google Acadêmico. Foram priorizados artigos de 2015 até 2020, levando em consideração a métrica utilizada de "Fator de Impacto de 5 anos" da Clarivate Analytics, conforme mostrado em Murziqin *et.al* (2020), com poucas exceções de trabalhos publicados antes desta data, apresentaram relação com o assunto deste artigo. Vale destacar que essas poucas exceções que foram utilizadas no presente trabalho, mesmo não estando dentro do período citado anteriormente, é devido ao fato de serem trabalhos com número de citações (>=50) em outros trabalhos. Recentemente, o assunto abordado neste trabalho tem ganhado maior atenção da comunidade científica. Por ser uma revisão com foco interdisciplinar, os autores não se preocuparam em fechar escopos no momento da pesquisa por publicações, onde os termos utilizados basicamente foram "*fake news*" e "*deepfake*" nos títulos dos artigos.

Inicialmente a pesquisa com o termo "*fake news*", dentro do período estabelecido, trouxe exatamente 5300 resultados, dos quais, seguindo o critério de número mínimo de citações, refinou-se para um total de 747 artigos. Em seguida, priorizou-se artigos com Qualis Capes no mínimo B2 como publicação, sejam periódicos ou eventos da academia, totalizando no final um total de 33 artigos utilizados neste trabalho.

Deve-se frisar que há algumas poucas exceções de trabalhos que não se adéquam ao requisito mínimo de citações (>=50 citações), devido ao ano de publicação do artigo ser relativamente recente (2019 e 2020), mas que atende ao requisito mínimo do Qualis Capes citado.

Em relação ao termo "*deep fake*", utilizando-se do mesmo critério de busca nos títulos dos artigos no período avaliado, foram encontrados 200 resultados. Buscando priorizar artigos com mais de 50 citações, chegou-se a um total de 9 artigos que atendiam a este critério. Contudo, levando em consideração que este é um termo muito recente, e atendendo as publicações com no mínimo B2 de classificação no Qualis Capes, optou-se por incluir artigos que não compreendiam o número mínimo de citações citado anteriormente, e refinou-se para um total de 5 artigos.

FAKE NEWS: CONCEITOS E CLASSIFICAÇÕES

Antes de serem apresentadas as definições atuais para o termo *fake news,* faz-se necessário tecer algumas considerações sobre o vocábulo "notícia". Notadamente, antes das primeiras menções à expressão "notícias falsas", a palavra notícia não carecia de um olhar semântico cuidadoso, tal qual é requerido neste texto. Tal demanda se apresenta a partir da consideração de que a junção destas duas palavras (notícia + falsa) configura uma figura de linguagem conhecida como oximoro, ou seja, o uso de dois termos com significados antagônicos que fornece uma nova ideia.

Historicamente, a palavra notícia sempre carregou consigo o conceito de verdade e fato, porém há poucos anos, quantos de nós cogitariam a possibilidade imediata de falsidade quando surpreendidos por um conhecido que iniciava sua fala a partir da frase: - Trago notícias! Ou mesmo ao assistir a um "noticiário" na TV. Não se trata aqui de afirmar que naqueles tempos a notícia sempre trazia verdades, mas de destacar a fragilidade que o termo alcançou nos tempos presentes.

Acerca do conceito de notícia, Silva (2009) diz que "Notícia é a socialização de quaisquer informações de caráter público, atual e singular e que atendem a diferentes interesses." No trabalho de De Souza Vieira e De Aquino (2020), as autoras apresentam o conceito de notícia, do ponto de vista de diversos autores. Além disso, elas buscam chegar a uma conclusão do conceito de notícia, através de um questionário respondido por 484 participantes. Em De Castro (2017), o autor busca analisar de que forma os algoritmos influenciam no conceito de notícia. Em ambos os trabalhos é possível extrair algumas conclusões a respeito da palavra notícia, a destacar: sua estreita relação com o jornalismo, seu caráter temporal que aponta para o presente e sua acepção de relevância no assunto tratado.

A notícia é o produto final da atividade do jornalista. Segundo Schudson (2003) o jornalismo deve cumprir um papel social que, amparado em valores éticos, conduz os cidadãos na direção de sua liberdade e autogovernança; deve noticiar com objetividade e precisão tendo consciência do poder e responsabilidade que a profissão encerra em si. É uma atividade que deve ser observada com bastante ponderação pela sociedade.

Fazer notícia, portanto, traz consigo uma série de responsabilidades que outrora se perderam em meio aos infindáveis entroncamentos das redes de computadores, do anonimato e das facilidades promovidas pelas redes sociais em se produzir e divulgar inverdades. Esse cenário favorece indivíduos mal-intencionados a passarem-se por jornalistas profissionais, travestindo suas mentiras como notícias. Desse modo, nos tempos atuais, foi preciso apropriar-se desse oximoro que será discutido a seguir, qual seja, a *fake news*.

Mas qual a diferença entre mentira e *fake news*? No trabalho de Mackenzie & Bhatt (2020), os autores buscaram definir o termo mentira. Para os autores, o mentiroso sabe que a informação declarada é falsa e tem a intenção expressa de enganar o alvo de forma a se obter alguma vantagem. O principal objetivo do mentiroso é o engano e que pode ser alcançado de forma engenhosa, fazendo com que as pessoas acreditem na informação falsa, seja por gestos, por meio de disfarces, por meio da ação ou até mesmo do silêncio.

Sobre *fake news*, ainda não existe um conceito formalmente definido e considerado um consenso no meio científico. É possível encontrar diversas definições para o termo *fake news* na literatura, conforme algumas apresentadas nos parágrafos a seguir.

Para Allcott & Gentzkow (2017) o termo *fake news* trata de textos noticiosos que são intencionalmente falsos e verificados como tal, e que tendem a enganar os leitores. No trabalho de Guess, Nyhan & Reifler (2018), os autores enfatizam a ideia de que *fake news* é uma forma de desinformar politicamente os leitores, com objetivo de obter lucros.

O termo *fake news* é atualmente compreendido como informação falsa que é utilizada pelos meios de comunicação para levantar suspeitas sobre notícias falsas que circulam, especialmente nas mídias sociais. A literatura classifica as sátiras, boatos e notícias fabricadas de igual modo, sendo informações que existem em uma quantidade relevante e amplamente divulgadas pelos meios de comunicação. (Tandoc, Lin & Ling, 2018).

Para Shu *et al.* (2017) para classificar as notícias como *fake news* é necessário que sejam apresentadas duas características básicas: a primeira é a não comprovação da veracidade dos fatos e a segunda é o propósito de iludir. Para os autores, as sátiras não são consideradas *fake news*, pois na narrativa explicita a falsidade através do humor e não tem a intenção de iludir. Os autores afirmam que as *fake news* são notícias fabricadas com informações falsas e divulgadas com o propósito de atender aos interesses de indivíduos ou grupos. Himma-Kadadas (2017) afirma que as *fake news* apresentam, em suas narrativas, uma roupagem jornalística e com elementos que possam gerar credibilidade.

Para Recuero & Gruzd (2019) as *fake news* apresentam três elementos essenciais para sua definição. O primeiro é o conjunto de elementos que compõem a narrativa com roupagens jornalísticas e noticiosas, o

segundo é a falsidade parcial ou total da narrativa e o terceiro é a intencionalidade de iludir ou gerar percepções equivocadas pela divulgação de tais informações.

Uma das definições mais utilizadas foi concebida por Claire Wardle para o Council of Europe (Wardle & Derakhshan, 2018) na qual a autora evita o uso do termo *fake news*, mas adota três categorias para a classificação da disfunção da informação: *misinformation*, *disinformation* e *mal-information*. *Misinformation* é quando a informação é falsa, mas a pessoa que a dissemina acredita que é verdadeira, logo, não possuindo intuito de causar danos a terceiros. Exemplos de *misinformation* são sátiras levadas a sério, dados e datas erradas, traduções mal realizadas, entre outras. *Disinformation* é quando o remetente, intencionalmente sabe que a informação propagada é falsa. Portanto, notícias falsas com o objetivo de prejudicar terceiros. *Mal-information* são notícias verdadeiras mas que, intencionalmente, buscam prejudicar o indivíduo, empresa ou país. Pode-se destacar vazamentos de informações ou mesmo discursos de ódio, que consistem em postagens realizadas nas mais diversas mídias, sejam impressas ou online, que possuem caráter polêmico, mobilizando um enorme número de indivíduos que passam a discutir e a expor seus diferentes pontos de vista que, por vezes, emanam o ódio relacionado à discordância de pensamentos e ideologias. Os responsáveis por priorizarem a violência através dos discursos de ódio são conhecidos como *haters*, ou também chamados de "odiadores" ou "*trolls*", conforme Rebs & Ernst (2017). Na Figura 1, a seguir, é possível ver a classificação dos tipos de disfunção, segundo Wardle & Derakhshan (2018).

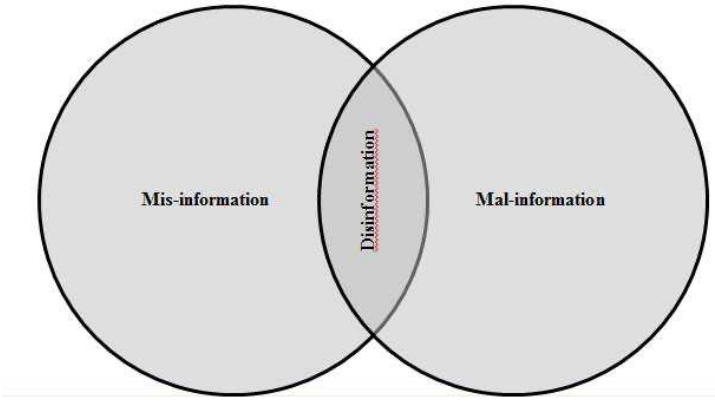

Figura 1. Tipos de disfunção.
Fonte: Elaborado pelos autores

Vale ressaltar que Wardle & Derakhshan (2018) fazem uma espécie de subclassificação das formas que essas disfunções podem ser apresentadas, conforme listadas no Quadro 1.

Quadro 1. Diferentes formas em que as disfunções são apresentadas

	Misinformation	***Disinformation***	***Mal-information***
Sátira ou paródia	√		
Falsas conexões	√		√
Conteúdo enganoso		√	
Falso Contexto	√		
Conteúdo impostor			√
Conteúdo manipulado		√	
Conteúdo fabricado		√	

Fonte: adaptado de Wardle & Derakhshan (2018)

Para melhor compreensão dos termos mostrados na Figura 1, alguns exemplos são ilustrados a seguir:
- *Misinformation*: Um ataque terrorista ocorrido em 20 de abril de 2017 na Champs Elysées em Paris, levou muitas pessoas em redes sociais a publicarem uma série de boatos, como por exemplo, a morte de um segundo policial. É possível perceber que a intenção da propagação, neste caso, é informar, sem o

objetivo de causar danos, buscando ser úteis. Porém, o efeito "câmara de eco"[1] ao propagar tal informação, sem checar a veracidade, pode causar transtornos até a verdadeira informação, de fato, vir à tona.[2]

- *Disinformation*: uma tentativa de manipulação nas eleições presidenciais da França de 2017, onde houve a criação de uma versão duplicada muito bem elaborada do jornal belga *Le Soir* com um falso artigo cujo conteúdo informava que o candidato à presidência Emmanuel Macron estava sendo financiado pela Arábia Saudita[3]. Outro exemplo foi a circulação de documentos online onde informavam que o mesmo candidato teria uma conta de banco nas Bahamas, país considerado paraíso fiscal[4]. É possível perceber, nestes dois exemplos, claras tentativas de propagação de conteúdo enganoso.

- *Mal-information*: Um exemplo notável foi durante a eleição nos Estados Unidos da América de 2016, sobre o uso de um e-mail particular de Hillary Clinton, fazendo referência ao período de sua atuação como Secretária de Estado americano, entre 2009 e 2013. Noticiar tal fato, às vésperas do pleito, mostrou que a intenção, por parte do candidato concorrente, de alguma forma, foi de causar danos à campanha de Hillary, obviamente não sendo o único fator com esse objetivo.

Todo conteúdo criado, independente do seu destino final, considerando a Internet como principal meio de propagação de mensagens, possui inumeráveis destinatários, pois qualquer mensagem pode ser propagada com facilidade para diversos destinos finais. Mas por trás de toda informação gerada, deve existir algum fator motivacional que sustente a sua criação. E no mundo das *fake news*, não é diferente. Utilizando-se a ideia das disfunções e suas subcategorias propostas por Wardle & Derakhshan (2018), o Quadro 2 mostra a relação entre as diferentes formas em que as disfunções são apresentadas e os objetivos da criação das *fake news*.

Quadro 2. Diferentes formas em que as disfunções são apresentadas

	Propaganda	**Influência Política**	**Lucro**	**Partidarismo**	**Emocionar**	**Provocar**	**Parodiar**	**Jornalismo pobre**
Sátira/ Paródia							√	
Falsas Conexões			√					√
Conteúdo Enganoso	√	√		√				
Falso Contexto								
Conteúdo Impostor	√		√			√	√	
Conteúdo Manipulado	√	√				√		
Conteúdo Fabricado	√	√	√		√	√		

Fonte: adaptado de Wardle & Derakhshan (2017)

Para facilitar a compreensão do leitor, abaixo conceitua-se cada uma das diferentes formas em que as disfunções são apresentadas.

- **Sátira ou paródia:** sem intenção de causar mal, mas com potencial para enganar;
- **Conteúdo enganoso:** uso enganoso de informação para enquadrar problema ou pessoa;
- **Conteúdo impostor:** quando fontes genuínas são representadas;
- **Conteúdo fabricado:** novo conteúdo é 100% falso, projetado para enganar e fazer mal;
- **Conexão falsa:** quando manchetes, imagens ou legendas não suportam o conteúdo;
- **Contexto falso:** quando conteúdo genuíno é compartilhado com informação contextual falsa;
- **Conteúdo manipulado:** quando informação ou imagens genuínas são manipuladas para enganar;

1 O efeito da câmara de eco, parte do princípio de que pessoas com os mesmos intereses ou pontos de vista interagem primariamente dentro do seu grupo, compartilhando informações que vão de encontro com as normas do grupo e tendendo a reforçar as mesmas crenças compartilhadas pelos membros.
2 Did London Muslims "celebrate" a terrorist attack on the Champs-Elysees? (2017, April 25). Retrieved October 20, 2020, from https://crosscheck.firstdraftnews.org/checked-french/london-muslims-celebrate-terrorist-attack-champs-elysees/
3 Was Macron's campaign for the French presidency financed by Saudi Arabia? (2017, March 21). Retrieved October 20, 2020, from https://crosscheck.firstdraftnews.org/checked-french/macrons-campaign-french-presidency-financed-saudi-arabia/
4 Did Emmanuel Macron open an offshore account ? (2017, May 08). Retrieved October 20, 2020, from https://crosscheck.firstdraftnews.org/checked-french/emmanuel-macron-open-offshore-account

Portanto, é possível perceber que as *fake news* possuem diversas nuances e detectá-las não é algo trivial, pois exige o uso de ferramentas, técnicas e outros artifícios que serão discutidos mais à frente.

Manipular textos para propagação de notícias falsas foi apenas o primeiro passo. Atualmente, um sofisticado meio de propagar notícias falsas que está ganhando força na internet são as *deepfakes*.

DEEPFAKE: UMA FAKE NEWS MAIS AGRESSIVA E ASSUSTADORA!!!

As *fake news* textuais continuam influenciando pessoas, seus pensamentos e comportamentos e, para piorar a situação, surgiu uma nova abordagem de *fake news* chamada *deepfakes*, que cresce e apresenta-se com um grau de dificuldade ainda maior para identificar o que é falso ou não.

O crescente aumento no número de dispositivos móveis em geral (smartphones e tablets), com poder de processamento cada vez maiores, diversos softwares, sites que permitem a postagem, compartilhamento de fotos e vídeos criam oportunidades para novas abordagens maliciosas para criação e propagação das *deepfake*s (Korshunov & Marcel, 2018).

Basicamente, as *deepfake*s utilizam técnicas de redes neurais, inteligência artificial e *deep learning* para manipulação de imagens e vídeos, alterando rostos, ou até mesmo vozes originais por outras. Conforme Lecun, Bengio & Hinton (2015), *deep learning* permite que modelos computacionais aprendam representações de dados com múltiplos níveis de abstrações. Graças a estes métodos, o estado da arte em diversos domínios tais como reconhecimento de voz e objetos, compreensão da linguagem natural, análises de sentimentos, entre vários outros, foi atualizado de forma significativa. Um exemplo de aplicativo desktop que permite criar *deepfake* é o FakeApp.

Vale destacar que, apesar de recentemente ter ganhado maior atenção por parte da comunidade acadêmica, a tentativa da troca de rostos não é novidade. Por volta de 1865, houve a primeira tentativa de trocar o rosto de alguém. Neste caso, foi a tentativa de colocar o rosto do então presidente americano Abraham Lincoln com o corpo do político John Calhoun (Güera; Delp, 2018).

Um dos maiores usos de *deepfake* é o *pornfake*, que, conforme Henry, Powell & Flynn (2018) conceituam, é uma técnica onde o rosto de uma pessoa é utilizado no lugar do rosto de outra em vídeos pornográficos. Normalmente o vídeo *fake* é disponibilizado em sites de pornografia (Floridi, 2018). Na Figura 2, mostra-se a imagem original (à esquerda) e a mesma após manipulação (à direita). É possível perceber que a expressão do rosto original não é a mesma da imagem manipulada.



Figura 2. Exemplo de manipulação com a técnica *deepfake*.
Fonte: Afchar *et al.* (2018)

Percebe-se que as *deepfakes* elevam as *fake news* textuais para um outro patamar, aumentando a dificuldade em serem analisadas e identificadas em reais ou não. E por ser um assunto que recentemente tem se tornado um *hot-topic*[5] na comunidade científica, há muito a ser investigado.

COMO AS NOTÍCIAS REAIS E *FAKE NEWS* SE ESPALHAM?

Segundo Parisier (2011) a estrutura da rede não é suficiente para que as *fake news* sejam propagadas. Para que a propagação aconteça é necessário a atuação ativa de dois elementos, na condução de tal processo. O primeiro deles configura-se na ação dos algoritmos de visibilidade que fazem parte das redes e o segundo

5 Assunto que está em evidência no momento

são as escolhas dos atores que decidem o que compartilhar, publicar ou tornar visível aos seus contatos virtuais. Conforme discutido anteriormente, toda notícia tem uma origem e um propósito. Teoricamente, uma notícia real antes de sair da fonte, é apurada, passando pelo processo de sua criação (escrita) até chegar a fonte disseminadora, tornando-se notícia real e confiável de fato. Devido a este processo é que nos jornais e telejornais, antes de uma informação ser divulgada e se tornar notícia, primeiramente tal informação deve passar pelo crivo de um ou mais editores, que seriam uma espécie de revisores, antes da sua publicação. Contudo, com as *fake news* esse processo é diferente, compreendendo três etapas, conforme apontam Wardle & Derakhshan (2018):

- **Criação**: Fase em que há a criação, de fato, da mensagem e apuradas com profundidade. Não valem mensagens recebidas de terceiros (outras fontes).
- **Re(produção)**: Quando a mensagem criada é trabalhada para ser transformada em um produto de mídia.
- **(Re)Distribuição**: Quando a mensagem é tornada pública, distribuída sem um destino final definido e sendo redistribuída "infinitamente".

Essas três etapas possuem total relação com os próximos três elementos que compõem o framework básico e envolvem todo o processo de propagação das *fake news*: Agente, Mensagem e Interpretador, conforme mostrado na Figura 3.

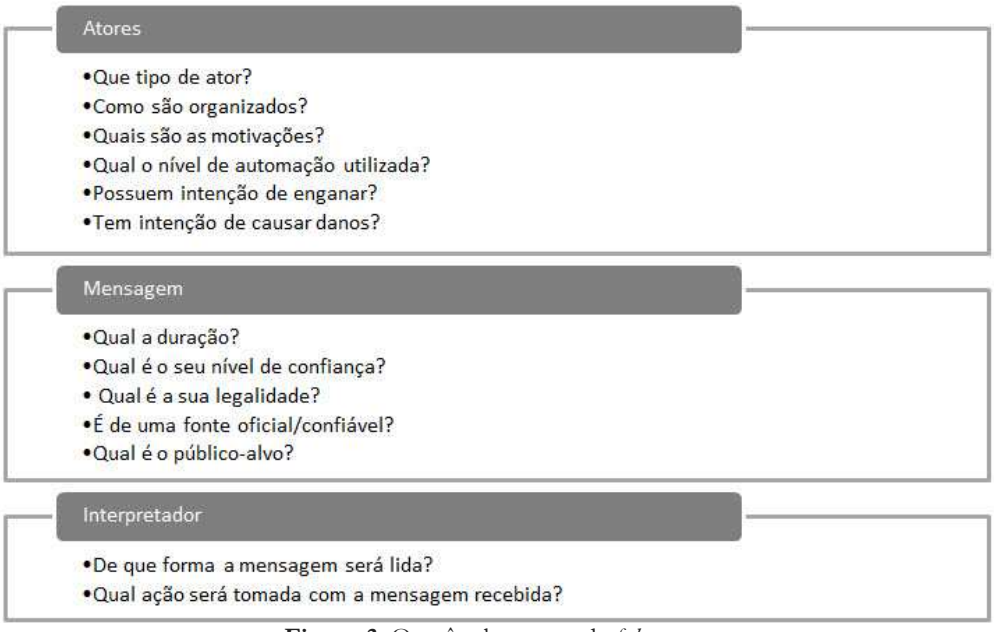

Figura 3. Os três elementos da *fake news*.
Fonte: adaptado de Wardle & Derakhshan (2018).

No trabalho de Passamani (2019), o autor apresenta, através de um diagrama, conforme exibido na Figura 4, as relações e interdependências existentes entre as *fake news* e notícias verdadeiras, desde as suas origens até a sua publicação, explicita-se claramente que o ato de publicar a notícia não quer dizer que é o fim do processo, onde o mesmo pode ser retroalimentado através da sua reprodução, nova distribuição, pelos diversos canais de propagação, como por exemplo as redes sociais, aplicativos de mensageria, sites, entre outros. As setas mostram o percurso criativo das notícias, sejam fakes ou não, e as setas pontilhadas representam como o leitor pode interferir na propagação dessas notícias.

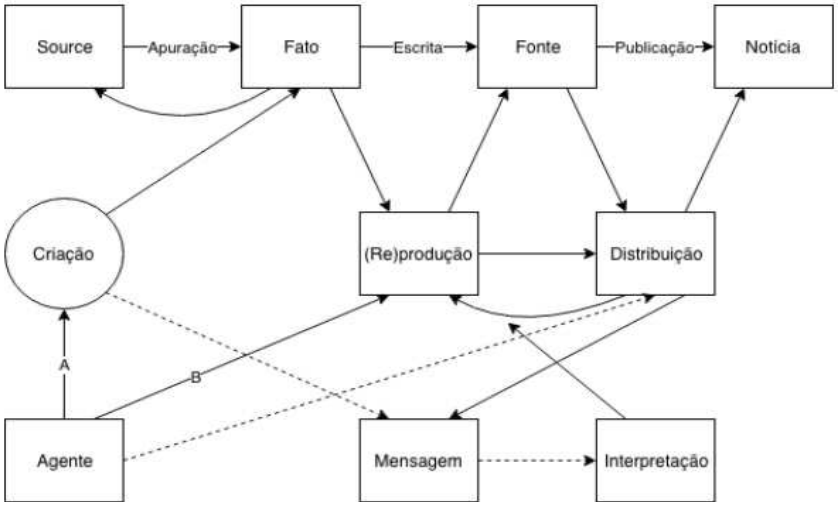

Figura 4. Relação entre notícias falsas e reais.
Fonte: Passamani (2019).

O termo *fake news* é relativamente recente (não confundir com a tentativa de manipular as pessoas com notícias/boatos/propagandas falsas) e, segundo Lazer (2018), cada vez mais tal nomenclatura se torna parte da vida diária das pessoas. Nesse contexto, é perceptível a queda da confiabilidade dos meios tradicionais de informação, tais como os jornais e televisão. Devido aos objetivos propostos através da propagação de *fake news*, várias ferramentas foram criadas com o intuito de classificar as notícias ou até mesmo as fontes e fatos que se apresentam com essas características. Diante dessa realidade torna-se importante saber conceituar notícias falsas, fatos e fontes.

Fatos são mensagens geradas em um instante pontual no tempo (*tweets*, entrevistas, debates, entre outros). A veracidade de qualquer informação deve ser apurada e atualmente existem diversas ferramentas que realizam essas apurações devolvendo a análise da mensagem e classificando o conteúdo em verdadeiro ou falso, seja totalidade ou parcialmente falso.

Notícias falsas são textos que buscam ter a característica e aparência verdadeiras, porém, possuem modificações realizadas de forma proposital com diversos intuitos, tais como: provocar emoção, angariar mais leitores ou enganar.

As fontes de informações falsas são sites ou mesmo pessoas que criam e/ou disseminam notícias ou fatos falsos. No trabalho de Subramanian (2017), observou-se que mais de 100 sites, fontes de *fake news*, relativos ao pleito eleitoral estadunidense de 2016, foram operados por adolescentes em uma pequena cidade da Macedônia. No trabalho de Sydell (2016), a autora destaca a existência de uma empresa chamada Disinfomedia, responsável por vários sites de *fake news*, com o objetivo principal de obter lucro e que possuía de 20 a 25 pessoas responsáveis pela criação e/ou disseminação das *fake news*. É possível ver no Quadro 3, os diferentes modos de difusão das *fake news*.

Quadro 3. Diferentes formas em que as disfunções são apresentadas

	Fatos	**Fontes**	**Notícias**
Explicação	Informações pontuais produzidas através de discursos ou de mensagens/vídeos em redes sociais ou programas de mensageria	Por onde a notícia é divulgada. O disseminador da notícia	Matérias com o intuito de informar os leitores
Alguém verifica?	Lupa, Fato ou Fake, Polifact (EUA), Snopes (EUA)	Décodex (francês)	Polifact (EUA), Snopes (EUA)
Como Verifica?	Através de investigações sobre a veracidade dos fatos apresentados	Existência de autores, parte editorial, informações gerais da notícia.	Confiabilidade dos autores

Fonte: adaptado de Passamani (2018)

No campo das redes sociais e sua relação com as *fake news*, o nível de propagação das notícias falsas pode ter maior alcance, tendo em vista o aumento no uso dessas plataformas conforme pode ser visto na Figura 5, que compara o aumento no número de usuários ativos entre os anos de 2018 e 2019, no Brasil.

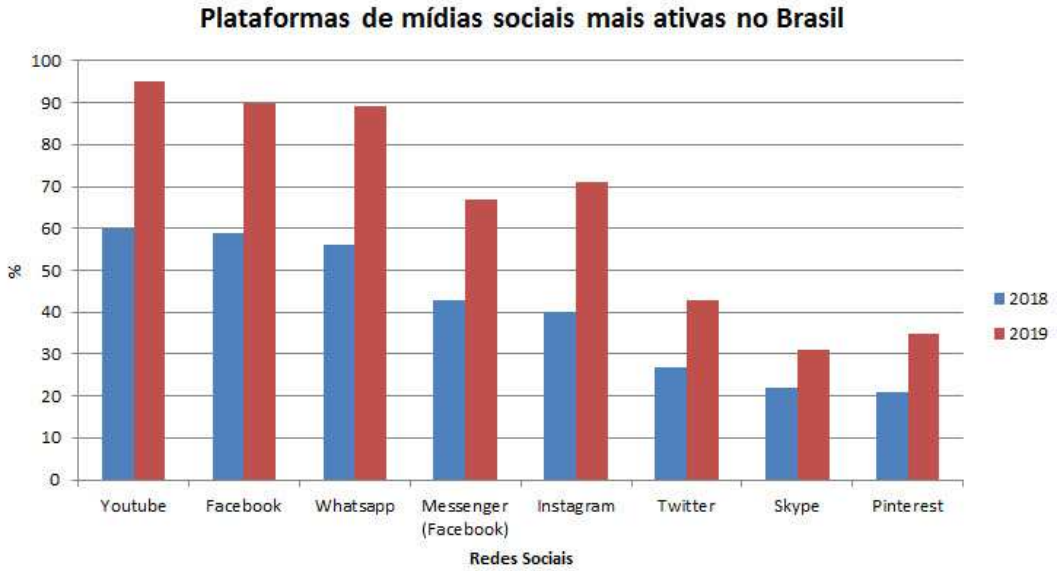

Figura 5. Comparativo de percentual de usuários ativos nas redes sociais.
Fonte: adaptado de Costa (2019)

De acordo com Ferreira (2018), o baixo custo para produção de conteúdo, sem se preocupar com a reputação de qualidade a longo prazo e a busca por rentabilidade em curto prazo, tendo em vista que o tempo de vida de sites de *fake news* é curto, somados ao formato das plataformas de redes sociais, que ainda não possuem ferramentas eficazes de combate às *fake news* que consiga avaliar se uma determinada notícia/informação é falsa ou não, são características ideais para propagação.

De acordo com Gabielkov (2016), 59% dos links compartilhados nas redes sociais não são sequer abertos por aqueles que compartilham, mostrando que as pessoas possuem mais disposição para compartilhar possíveis *fake news*, do que lê-las de forma detalhada buscando a veracidade, tendo como base a sua capacidade crítica em um simples resumo da informação ou até mesmo em um resumo do resumo, onde percebe-se claramente o comportamento típico do usuário de hoje: a capacidade de atenção é cada vez menor.

A pesquisa realizada por Bakshy, Messing & Adamic (2015), mostra que grupos de amigos no Facebook, de forma geral, são agrupados por ideias e opiniões em comum. Cada grupo busca propagar informações que estão em consonância com a linha de seus pensamentos, ideologias, mesmo sabendo que aquela informação compartilhada é uma *fake news*, criando o que Parisier (2011) chama de "filtros-bolha", onde a opinião contrária não tem vez. Portanto, as *fake news* possuem as mesmas características de reverberação de "câmaras de eco" digitais, apresentado por Sustein (2001).

Uma outra forma para conseguir a propagação de *fake news* de maneira automatizada, rápida e capaz de alcançar um maior quantitativo de pessoas nas redes sociais, é através do uso de *bots* (robôs), conforme Messias *et al.* (2013). No trabalho de Watts (2017), o autor estima que em torno de 60 milhões de contas no Facebook são de robôs. Para Varol (2017), entre 9% a 15% das contas existentes no Twitter são de robôs. Basicamente os *bots* são programas para realizar tarefas repetitivas em busca de obtenção de dados e pode-se dizer que esses dados hoje tornaram-se mais um tipo de *commodity*. Eles podem monitorar diversas variáveis por parte do usuário, tais como onde o usuário clicou, quais páginas navegou, se o usuário compartilhou alguma informação, dessa forma moldando o perfil do usuário de acordo com o seu gosto, enviando conteúdo de encontro com o seu perfil.

Segundo o trabalho de Shao *et al.* (2017), uma parcela pequena das contas no Twitter influenciam de forma significativa o tráfego de *fake news* dentro da rede social citada.

> Essas contas são provavelmente *bots*, e descobrimos várias estratégias de manipulação que usam. Primeiro, os *bots* são particularmente ativos em amplificar notícias falsas no estágio inicial de

> propagação, antes que o conteúdo seja viral. Em segundo, os *bots* miram usuários influentes através de respostas e menções (Shao *et al.*, 2017, p. 11).

A fala de Shao *et al.*, mostra que a velocidade e o volume de *fake news* propagadas, evidencia a presença de grupos que controlam perfis, que são chamados de perfis *fake*, utilizados para disseminação de *fake news*, através do uso de *bots*. Portanto, é possível perceber que cada vez mais as técnicas para propagação de *fake news* estão ficando mais sofisticadas e que as diversas empresas, detentoras das plataformas de mídias sociais, terão que criar mecanismos para acabar ou minimizar as *fake news* e consequentemente os efeitos nocivos que causados a sociedade.

AS RELAÇÕES ENTRE *FAKE NEWS* E AS DIFERENTES ÁREAS DO CONHECIMENTO

Nesta seção do artigo, serão apresentados ao leitor trabalhos relevantes que foram publicados em eventos ou periódicos que atenderam aos critérios definidos na seção de metodologia. Os trabalhos têm relação com as *fake news* e diversas áreas de conhecimento, como por exemplo, a computação, matemática, a neurociência, a psicologia, entre outras. O objetivo da seção é entregar ao leitor um panorama de como o assunto *fake news* e, recentemente, as *deepfakes* estão "navegando" pelas mais diversas áreas do conhecimento, mostrando que é um assunto atual, interdisciplinar e que pode dar alguns "*insights*" para novas discussões.

Fake news, *Deepfake* e a Computação

Muitas são as investigações sobre a relação entre *fake news*, *deepfakes* e a computação. De fato, há uma relação muito rica envolvendo estes termos e áreas. As pesquisas apontam diversas abordagens, permeando, desde redes neurais até a utilização da clássica Lei de Benford (Tam Cho & Gaines, 2007), *fake news* e a *deepfake*.

No trabalho de Kong *et al.* (2020), os autores utilizaram técnicas de processamento de linguagem natural para analisarem textos e melhorarem modelos de *deep learning* (Lecun, Bengio & Hinton, 2015). No artigo de Ruchansky, Seo & Liu (2017), os autores criaram um modelo para buscarem uma predição automática e mais eficiente de *fake news*. O modelo combina três características existentes nas *fake news* para realizar as buscas e detectá-las. As características são: o texto de uma notícia; o feedback do usuário que recebe a notícia; e a fonte que propaga a notícia.

Existem outras abordagens envolvendo a computação em busca da solução ou minimização dos impactos das *fake news* na sociedade. Uma dessas técnicas é o uso de *Blockchain* (Zheng *et al.*, 2018) que, inicialmente, é muito conhecido devido a sua adoção no mundo das moedas virtuais, onde pode ser citado como exemplo mais clássico o Bitcoin (Swan, 2015). No trabalho de Shang *et al.* (2018), os autores buscaram rastrear a origem das *fake news*, levando em consideração o armazenamento distribuído, descentralização e outros recursos existentes no *Blockchain*. No trabalho de Shae & Tsai (2019), os autores deixaram claro que para atacar as *fake news*, é necessário um esforço interdisciplinar, devido a isso, eles propuseram uma plataforma *Blockchain* baseada em Inteligência Artificial que permite colaboração entre os pesquisadores na área de *Blockchain*, Inteligência Artificial e jornalistas das mais diversas mídias trabalharem em conjunto contra as *fake news*.

Outra técnica que tem investido contra as *fake news* é a *machine learning*. Existem diversos conceitos acerca de *machine learning*, conhecido em português como aprendizado de máquina. Um dos conceitos é que *machine learning* pode ser um sistema capaz de modificar seu comportamento de maneira autônoma, tendo como base a sua própria experiência. A interferência humana é mínima (Alpaydin, 2020). No trabalho de Ahmed, Traore & Saad (2017), através de técnicas de *machine learning*, os autores propuseram um modelo de detecção de *fake news*. No trabalho de Aphiwongsophon & Chongstitvatana (2018), os autores utilizaram de três técnicas de *machine learning* para detectar *fake news*, realizaram comparações entre elas e para identificarem qual apresentou maior percentual de eficácia na detecção. Os resultados mostraram que duas técnicas utilizadas alcançaram um percentual de 99,90% de eficácia.

No trabalho de Tschiatschek *et al.* (2018), os autores buscaram evitar a redução de *fake news* na rede. Para isso, eles criaram um algoritmo chamado *Detective* que utiliza de inferências bayesianas. Os resultados alcançados apontaram que a abordagem do algoritmo se apresenta bastante competitiva em comparação com outros utilizados e que tem funcionalidade semelhante.

No recente trabalho de Parikh *et al.* (2019), foi proposto, um modelo para calcular o impacto de uma *fake news*. Os autores levaram em consideração diversos fatores para realizar essa pontuação, tais como o escopo da notícia, a reputação do site que a publica e a popularidade de quem a propaga.

No trabalho de Korshunov & Marcel (2018), os autores mostraram como tem sido cada vez mais fácil criar *deepfakes* usando as *Generative Adversarial Network* (GAN) (Ledig *et al.*, 2017). Os resultados apontaram que os vídeos criados usando a técnica GAN são mais difíceis de serem detectados pelos sistemas de reconhecimento de face e métodos de detecção existentes. Em Hall (2020), os autores buscaram investigar se existe alguma relação entre a Lei de Benford e os *bots* de redes sociais. Os resultados evidenciaram que a aplicação da Lei de Benford ajuda na detecção de contas online maliciosas automatizadas e as suas atividades nas redes sociais.

É possível perceber que a comunidade científica tem se esforçado em utilizar diferentes abordagens em busca da solução definitiva, em relação às *fake news* e *deepfakes*, porém, percebe-se que a solução definitiva do problema ainda está longe de ser alcançada. Nesse sentido, a neurociência e a psicologia cognitiva oferecem ricas contribuições como as que seguem.

Fake News e deepfake sob a ótica do cérebro e do comportamento humano (da psicologia cognitiva)

Depois de uma abordagem focando nas ciências exatas, matemática e computação, que mostraram as relações existentes das *fake news* e *deepfakes*, nesta seção, abordam-se as relações dos termos considerando a psicologia e a neurociência.

A neurociência envolve diversas áreas, portanto, é multidisciplinar e tem como núcleo de estudo o sistema nervoso e diversos aspectos relacionando-se com a cognição (Pessoa, Botinha & Costa, 2018). Na psicologia, a neurociência tem como objetivo principal a análise do comportamento social e suas reações perante as *fake news* e *deepfakes* (Pennycook & Rand, 2020).

Segundo Recuero & Gruzd (2019), os envolvidos em redes sociais apresentam uma tendência em compartilharem informações baseadas em suas próprias crenças e percepções, especialmente em contextos que geram desconfortos com embates através de discussões ou polêmicas. Nesse aspecto a mídia social comporta-se de maneira propensa a apresentar redes de conversações extremamente polarizadas, constituindo-se por polos opostos.

Segundo Filho & Calabrez (2017), o processo de tomada de decisões tem como estrutura chave o córtex pré-frontal. O córtex pré-frontal regula uma série de funções e, uma delas é o processo de tomada de decisões, no planejamento e responsabilidade no processo de atualização profissional. O córtex pré-frontal é inibido com o uso de certas substâncias, por exemplo o álcool.

O ser humano tem uma tendência em criar mecanismos de fuga para algumas decisões que gerem desconforto. Nesse sentido é interessante refletir sobre o que é mais fácil: começar uma dieta agora ou deixar para a próxima semana? Com certeza deixar para a próxima semana é mais confortável para o cérebro, que faz ajustes mudando o comportamento e preservando a crença.

Quando é difícil agir de fato ou somos ludibriados pela emoção, o ser humano também pode apresentar uma tendência natural à inércia. O cérebro faz um processo de autoengano, onde muda a crença para não desestabilizar o indivíduo. Por exemplo, um casal com idade próxima de 18 anos, recém ingressados na faculdade, com 3 meses de namoro, um pergunta para o outro da seguinte maneira: "amor você me ama?" O outro responde "sim". O primeiro complementa: "para sempre?" O outro responde "para sempre". Estatisticamente, para um casal recém-admitido na faculdade, é pouco provável que, no mundo em que estamos vivendo permaneçam juntos para sempre. Mas por que eles respondem que é "para sempre?" No cérebro de alguém que está apaixonado, na "pastinha namoro" só tem coisas bonitas e acrescentar uma informação que desarmoniza não é agradável.

Nesse sentido é importante imaginar por que uma pessoa que é considerada bem informada, acredita piamente que uma notícia falsa é verdadeira. Tal posicionamento leva a reflexão sobre as bolhas e os efeitos cascatas que ocorrem no uso das redes sociais. Os algoritmos das redes identificam e propagam as informações de acordo com o interesse dos grupos e dos usuários. Os usuários recebem informações que deixam o cérebro com a sensação de estabilidade. Assim, torna-se cômodo ao usuário acreditar que a informação que chega é verídica (ou não) de acordo com seus interesses. O mesmo ocorre com o efeito cascata, ou seja, o usuário decide compartilhar, publicar ou deixar visível alguma informação que não lhe causa nenhum desconforto. Os usuários que vivem em um polo, dificilmente vão mudar o comportamento e serem críticos às informações que recebem. Tem uma charge que o senhor questiona o jovem usando a seguinte frase: "como essa notícia pode ser falsa se ela é exatamente o que eu penso?"

Existem trabalhos na literatura que utilizam de variáveis, tais como, batimentos cardíacos ou movimento dos olhos para analisar se um determinado conteúdo é *fake news* ou não. Este é o caso do

trabalho de Lutz *et al.* (2020), onde os autores realizaram um experimento que envolveu 42 assuntos diferentes, analisando as variáveis citadas anteriormente. Os autores afirmaram que a classificação de uma notícia como *fake* ou não, não é puramente cognitivo, mas também envolve processamento da informação afetiva.

Uma das habilidades dos seres humanos é a de interagir socialmente, construindo relações e conhecendo o mundo que os rodeia. O crescimento de rostos gerados por computador tem intrigado autores a questionarem até que ponto os seres humanos conseguirão, através de uma tela, reconhecer uma face real ou gerada por computador. No trabalho de Manohar (2020), o autor mostra que ainda é necessária maior investigação nesta área, com o objetivo de capacitar os seres humanos a saberem diferenciar rostos gerados por computador e rostos humanos reais.

Com o propósito de melhor compreensão de alguns processos psicológicos que podem nortear as discussões, diante das diversas áreas da psicologia, reconhecendo o valor de cada uma delas, elegeu-se a psicologia cognitiva como aquela que pode fornecer informações mais significativas para os propósitos desse estudo, uma vez que se busca teorias que ofereçam um suporte para comportamentos humanos que são percebidos e afetados de forma generalizada.

Inicialmente, buscou-se um suporte nos trabalhos escritos pelo psicólogo israelense Daniel Kahneman, especialmente em sua obra "Rápido e Devagar: Duas formas de pensar"**,** que possibilitou um prelúdio pertinente para essa discussão. De maneira didática, o autor aponta duas formas que nosso cérebro recebe e lida com as informações recebidas. Para isso, nomeia-as de "Sistema 1" e "Sistema 2", que de maneira resumida, está apresentado a seguir.

Para lidar de maneira sadia com o excesso de informações que se recebe o tempo todo, nossa evolução biológica tratou de criar um sistema de filtros, que julga a complexidade e a importância dos dados recebidos, a fim de, para cada situação, proporcionar respostas automatizadas ou mais elaboradas cognitivamente. Tal mecanismo se faz necessário, uma vez que se fosse preciso dispender muita atenção para beber um copo d'água, por exemplo, possivelmente haveria uma sobrecarga do nosso cérebro. Para resolver esse problema, a automatização de processos foi a solução encontrada por nosso aparato neuro cognitivo. Assim, atividades que são realizadas com frequência são incapazes de promover uma sobrecarga cognitiva, sendo, inclusive, frequentemente realizadas em concomitância com outras atividades de mesma natureza. Também é comum o fato do nosso cérebro realizar uma tarefa automatizada com uma atividade de grande demanda cognitiva, por exemplo, dirigir para aqueles com experiência ao volante e pensar em uma questão importante do trabalho (Kahneman, 2011).

A fim de diferenciar esses dois modos de operação, Kahneman (2011) nomeou os processos automatizados de Sistema 1 e, os processos de maior demanda cognitiva, de Sistema 2. Para o leitor é apenas necessário compreender que o primeiro sistema nos mantém em um confortável estado cognitivo. O sistema 1 percebe e resolve problemas com velocidade, fazendo uso das informações disponíveis e que foram arquivadas a partir de nossas experiências. O sistema 2, do pensamento complexo, de atenção e concentração, somente é acionado quando o sistema 1 não é capaz de proporcionar uma solução satisfatória para a situação enfrentada ou, quando voluntariamente o indivíduo escolhe acioná-lo. No entanto, o sistema 1 frequentemente se engana perante tarefas simples ou capciosamente elaboradas para que lhe parecem simples. Dessa forma, o sistema 2 não entra em operação, pois aparentemente o sistema 1 resolveu o problema. É preciso destacar que a solicitação do segundo sistema implica em sair dessa zona de conforto cognitiva e despender tempo, atenção e energia para a resolução de uma demanda específica, trabalho esse que, frequentemente, será evitado.

A maioria das teorias apresentadas a seguir e destacadas no Quadro 5, tem origens da relação estabelecida entre essas duas formas de pensar. Além disso, conforme discutido, a internet é principal meio de propagação das *fake news,* muitos processos psicológicos destacados são reforçados por algoritmos presentes em sites de busca e redes sociais.

A primeira teoria em destaque é conhecida por "Efeito Halo", definida por Kahneman (2011, p.74) como "A tendência de gostar (ou desgostar) de tudo que diz respeito a uma pessoa – incluindo coisas que você não observou". Poderia ser também abordado como um fator de atratividade (Huang, Zhu, Mustafaraj, 2019). Este é um viés de comportamento bastante significativo e influenciador do sistema 1, que opera a todo momento. Sua influência é abrangente pois, mesmo que se tenha consciência dele, ainda assim pode-se cometer erros de análise e interpretação. A teoria do efeito Halo informa que a primeira impressão que se tem sobre uma pessoa, ainda que possa ser modificada, tem um maior peso na análise futura. A título de exemplificação, foi adaptado, um exemplo, de Kahneman (2011). Assim, consideram-se dois sujeitos e suas características conforme Quadro 4:

Quadro 4. Características dos indivíduos.

	João	**Miguel**
Características	Inteligente	Invejoso
	Dedicado	Impulsivo
	Impulsivo	Crítico
	Crítico	Determinado
	Determinado	Dedicado
	Invejoso	Inteligente

Fonte: adaptado de Kahneman (2011)

Nestes casos, a maioria das pessoas seria mais favorável ao João do que ao Miguel. Isso acontece pela tendência que todo ser humano tem em dar um maior peso as características iniciais das pessoas que lhes são apresentadas. No entanto, o efeito Halo tem uma presença bastante veloz no julgamento onde, muitas vezes, não se precisa da exposição a tantas características para determinar uma relação de (des)apego com uma pessoa. O efeito Halo pode explicar porque tantos nortes-americanos acreditaram em Donald Trump, quando o na época presidente dos EUA, sugeriu que tomar desinfetante seria uma maneira eficiente de eliminar/prevenir a Covid-19. É desnecessário dizer que o resultado dessa crença foi um desastre para as pessoas que puseram em prática as palavras do ex-presidente.

Com um mecanismo similar ao efeito Halo, Slovic *et al.* (2002) apud Kahneman (2011) desenvolveu um conceito chamado de "Heurística do Afeto"**,** onde "as pessoas deixavam que suas simpatias e antipatias determinassem suas crenças acerca do mundo" (p.99). Em outras palavras, as atitudes emocionais das pessoas diante de coisas, notícias ou eventos tem uma influência determinante em suas crenças. Neste caso, uma vez que a pessoa tome consciência desse efeito (acesso ao sistema 2), suas influências podem ser contornadas com facilidade. Apesar disso, a heurística do afeto é um dos padrões psicológicos que será amplificado pelos algoritmos da internet, configurados para mapear os rastros da navegação dos usuários.

A "Ilusão de veracidade" é alimentada pelos algoritmos das redes sociais e sites de navegação. Segundo Kahnemam (2011) esse fenômeno está associado a uma sensação de conforto cognitivo, manifestada por uma ausência de ameaças, novidades ou necessidade de mobilizar esforços para dar atenção a algum evento. Esse mecanismo manifesta-se em contextos onde o indivíduo é exposto a uma variedade de informações verdadeiras onde, propositalmente, é implantada uma informação falsa. Nestes casos, a Ilusão de veracidade é um recurso bastante utilizado por alguns sites com claras posições políticas e que propagam *fake news*. Alguns desses sites transmitem notícias verdadeiras quando estas estão associadas a uma boa imagem de seus políticos, no entanto, colocam notícias falsas que comprometem a imagem dos adversários políticos.

Conforme observado, o campo político é um dos mais influenciados pela divulgação das *fake news*. Nesse contexto, destaca-se um mecanismo psicológico conhecido por "Heurística da disponibilidade", compreendida como "o processo de julgar a frequência segundo a facilidade com que as ocorrências vêm à mente" (Kahneman, 2011), ou seja, o cérebro humano apresenta uma tendência de atribuir maior presença estatística a um evento, se este lhe aparece com mais facilidade na memória. Por esse motivo, por exemplo, existe uma predisposição das pessoas aceitarem que as celebridades se divorciam com maior frequência que a média populacional, haja vista suas vidas estarem mais expostas na mídia. Sob o esteio da Heurística da disponibilidade uma *fake news* de grande impacto e crença por parte dos leitores, pode se tornar um evento intenso em suas memórias; como foi discutido, quando se atinge esse ponto, tal evento toma um sentido que erroneamente tem grande frequência estatística na mídia.

Apoiando-se na Heurística da disponibilidade, Kuran & Sunstein (1998) propuseram o conceito de "Cascata de disponibilidade", trazendo a ideia de um ciclo de eventos autossustentáveis que se inicia por meio de uma notícia sem grande importância, mas que ganha enorme notoriedade do público, gerando um processo de retroalimentação a partir da cobertura repetitiva da mídia. Ao trazer essa ideia para o campo das *fake news,* fica perceptível a presença dessa teoria, notadamente em grupos de aplicativos de mensagens, que, nesse caso, cumprem um papel semelhante ao da mídia no processo de retroalimentação da *fake news*. Assim como uma notícia sem importância pode ser acelerada intencionalmente, por exemplo, por empresários com interesses no mercado financeiro, uma *fake news* pode ser impulsionada por organizações criminosas, com interesses especialmente direcionados a influenciar politicamente a opinião as pessoas. A cascata de disponibilidade tem um dispositivo de funcionamento bem semelhante ao "efeito câmara de eco", frequentemente mencionado em outros estudos sobre *fake news* (Zhou & Zafarani, 2020; Lazer, 2018; Grinberg *et al.* 2019).

Os aplicativos de mensagens, bem como as redes sociais, possuem um papel fundamental em um viés psicológico conhecido como "Efeito Bandwagon". Inicialmente o Efeito Bandwagon foi descrito pelo pesquisador Lieberstein em 1950, que relatou a influência social na aceitação de ideias pelas pessoas, mesmo que isso signifique ir contra as ideias pessoais para que o indivíduo se mantenha alinhado às ideias de seu grupo de pares (Colman, 2015). Dessa forma, além da internet, a chamada Heurística de Bandwagon, ou ainda, Efeito adesão, possui grande influência nos círculos sociais presenciais, como reuniões familiares e de amigos. Nos últimos anos, discussões com teor político têm sido tema frequente nesses encontros; não é raro que nessas reuniões haja tentativas de convencimento de seus pares, muitas vezes, tendo como escopo de argumentação as notícias falsas. Vale destacar que o Efeito adesão tem pouca ou nenhuma relação com a dinâmica de operação dos sistemas 1 e 2 discutidos no início desta seção.

Na esteira dos mecanismos psicológicos com grande amplificação pelos algoritmos de redes sociais e mecanismo de busca, destaca-se também o "Efeito de validade" (Boehm, 1994) que condiciona os usuários das redes a acreditarem em uma *fake news,* na medida em que esta notícia falsa é repetida muitas vezes. Um exemplo é citado por Joseph Goebels, que afirmava que uma mentira repetida mil vezes torna-se verdade, o Efeito de validade tem desempenhado um grande papel na manipulação da população. Algumas são repetidas tantas vezes que, mesmo quando desmentidas por fontes oficiais, ainda existe resistência quanto à sua falsidade por alguns usuários. Este viés também é apontado por outros autores como PennyCook, Cannon & Rand (2018), que pontuam que, por menor que seja a exposição anterior a uma notícia, o indivíduo pode ser passível de considerar verdadeira uma informação, após observar a repetição da mesma. Outro termo encontrado na literatura para esse fenômeno recebe o nome de "Efeito da Verdade Ilusória" (Safieddine & Ibrahim, 2020). No trabalho de Liv & Greenbaum (2020), os autores afirmaram que as *deepfakes* não criam somente representações da realidade, mas podem ser utilizadas para criarem falsas memórias.

O "Viés de conservadorismo" (Basu, 1997) é descrito como um comportamento que os indivíduos adotam quando são confrontados sobre suas crenças e ideias, de maneira a posicionar-se firmemente em sua defesa. Para Kahneman (2011) é um comportamento que ajuda a manter a estabilidade das nossas vidas, mantendo-a "coesa junto ao ponto de referência" (p. 329). Pode-se interpretar esse viés cognitivo como uma característica moldada pela natureza, que vai ao encontro de nossos mecanismos de sobrevivência; os algoritmos aqui não possuem grande influência, mas é possível acionar o referido sistema 2, a fim de provocar autorreflexão. Apoiando-se no que foi discutido, observa-se que o viés de conservadorismo contribui para a formação de muita resistência às *fake news,* especialmente daquelas referentes a assuntos polêmicos e de ordem moral. No Quadro 5 é apresentado o quanto os algoritmos podem influenciar psicologicamente as pessoas.

Quadro 5. Vieses psicológicos e influências dos algoritmos em seu reforço.

Grau de influência de algoritmos	Viés psicológico
Grande	Ilusão de Veracidade Heurística da Disponibilidade Cascata de Disponibilidade Efeito de Validade
Moderado	Heurística do afeto Efeito Bandwagon
Baixo	Viés de Conservadorismo Efeito Halo

Fonte: Dos próprios autores

É relevante atentar como as relações sociais e as escolhas do indivíduo são afetados por algoritmos que, despidos de sentimentos, manipulam os comportamentos e as relações interpessoais, sejam elas virtuais ou físicas, onde as pessoas não conseguem ou negam-se a diferenciar o que é real daquilo que não é real. Dentro do ciberespaço (Castells, 2002) e das ciberculturas (Pierre, 1999), bolhas criadas, como por exemplo, as polarizações políticas, criadas pelas *fake news* e *deepfakes* não serão alteradas, enquanto os algoritmos continuarem a operar a partir de objetivos obscuros que atendam aos interesses de determinados grupos ou indivíduos.

Portanto, faz-se necessário ampliar as discussões de forma a incluir o papel que desempenham as grandes empresas de tecnologia que atuam na internet. Sua responsabilização não pode mais ser desconsiderada como ocorreu até o momento, pois os efeitos deletérios que seus algoritmos, somados aos mecanismos psicológicos apresentados, constituem um grande desafio para o controle de notícias falsas.

FAKE NEWS, DEEPFAKE E A POLÍTICA: SEUS IMPACTOS NAS POLARIZAÇÕES

Uma das áreas em que as *fake news* e *deepfakes* se impõem como estratégias para manipulação de opiniões e conhecimento são as ciências políticas. Principalmente a partir das eleições americanas de 2016 (Grinberg, 2019), onde se discute, por exemplo, se aconteceu ou não interferências da Rússia na vitória do então candidato Donald Trump.

No aspecto político, cada vez mais observa-se uma polarização partidária muito intensa, dividida em dois vieses. No Brasil, claramente percebe-se as duas posições mais fortes: esquerda e direita, assim como nos EUA ou se é republicano ou democrata. E as *fake news* e *deepfakes* tem contribuído significativamente no aumento dessa polarização.

No trabalho de Dias & Oliveira (2019), os autores, através de uma pesquisa bibliográfica, compararam a imprensa tradicional e as redes sociais digitais, analisaram o monopólio da verdade e seu impacto sobre o eleitor, que altera resultados eleitorais. Os autores concluíram que os políticos utilizam a mentira como um instrumento em favor de suas atuações políticas, que podem trazer prejuízos significativos à democracia e aumento da polarização.

No artigo de De Carvalho (2020), foi realizada uma revisão de literatura que relaciona a produção das *fake news* com incentivos econômicos e interesses políticos. No trabalho de Nascimento (2020), para a compreensão das *fake news*, o autor utiliza a perspectiva da pensadora judia Hanna Arendt, que viveu o regime totalitário na Alemanha do século XX, relacionando-a com o conceito da mentira organizada. O trabalho identifica que as *fake news* são uma prática da propaganda de massa, que foi originada de uma ideia totalitária. O autor conclui que a educação, através de seus atores, professores, professoras, estudantes é o único modo de resistência à mentira organizada e às *fake news*, tendo a responsabilidade de agir e resistir contra o cenário criado pelas *fake news*.

Vale destacar também que tem se buscado uma proliferação de esforços para criação de rodas de discussão, palestras e apresentações sobre *fake news*, tanto no ensino fundamental quanto no ensino médio, conforme o trabalho de Jones (2017). Porém, é incerto ainda se tais esforços melhoram as avaliações da credibilidade da informação ou se tais efeitos persistirão ao longo do tempo, tanto no sentido de evitar com que as pessoas sejam vítimas das *fake news* ou em mostrar os efeitos que sua produção pode causar.

Em Lelo (2020), o autor analisou o nível de confiança que os eleitores brasileiros atribuem a boatos e checagens fornecidas por empresas especializadas em *fake news*, foi utilizado como referência o atentado, ao então candidato à presidência nas eleições de 2018, Jair Bolsonaro. O autor realizou uma pesquisa, através de um questionário online com 108 respondentes, onde os resultados obtidos confirmaram alguns argumentos de pesquisas anteriores tais como: (a) os cidadãos utilizaram-se de boatos para se informarem; (b) as preferências políticas influenciaram a credibilidade ou ceticismo diante de *fake news*.

No trabalho de Castillo, Santos & Castro (2020), foi feito um olhar sobre as *fake news* envolvendo a COVID-19 sobre a perspectiva política. Os autores utilizaram de métodos quantitativos e qualitativos, realizando análises estatísticas e análises de conteúdo sobre a temática política, COVID-19 e *fake news*. Os autores chegaram à conclusão que a disseminação das *fake news* em torno da COVID-19 não é diferente da disseminação de outras notícias falsas.

Na abordagem do trabalho de Oeiras, Leite & Castro (2018), realizou-se uma revisão de literatura, analisando o papel da mídia na construção social da realidade. Para sua realização, foram cruzados os trabalhos de autores que trataram os movimentos sociais com o objeto de estudo e sua relação com as *fake news* que circularam na internet no período pós-*Impeachment* de 2016 no Brasil. Os autores concluíram que a procura pela credibilidade e a disseminação de *fake news* contra a sua posição ideológica, abriu um espaço de debate da sociedade em rede, sobretudo em discussões de cunho político.

Percebe-se que as *fake news*, no contexto político, possuem um impacto significativo sobre a opinião da sociedade, como relatado no trabalho de Van der Linder, Panagopoulos & Roozenbeek (2020), onde é analisado por um viés psicológico do efeito das *fake news* nos liberais e conservadores e a relação das fontes de mídia com o termo *fake news*. Como mencionado anteriormente, percebe-se um aumento na polarização partidária nas redes sociais, e certamente, muito influenciada pelas *fake news*, especialmente em períodos eleitorais, como os de 2016 nos EUA, 2018 no pleito presidencial no Brasil e em 2020 nas eleições presidenciais norte-americanas e municipais no Brasil. No Brasil, o Tribunal Superior Eleitoral, através de propagandas nas mídias, buscou chamar a atenção dos eleitores para prestarem muita atenção às *fake news* e *deepfakes*, empenhando-se de alguma forma em sensibilizar os eleitores sobre os impactos negativos que tais abordagens podem causar.

CONSIDERAÇÕES FINAIS

O presente artigo buscou contribuir junto à comunidade científica, com uma revisão de literatura com caráter interdisciplinar envolvendo as *fake news* e *deepfakes*. Até o presente momento, este é o primeiro trabalho em português com essa abordagem.

O ato de produzir notícias falsas, ou o seu acrônimo em inglês, *fake news*, é datado de um longo tempo. Porém, recentemente tem ganhado maior notoriedade e atenção da sociedade global, e, obviamente, dos pesquisadores da academia, em relação aos impactos causados na população, tanto nos aspectos das relações sociais e, mesmo, no impacto na saúde mental. Porém, com a evolução da tecnologia, novos recursos surgiram e com eles surgiram as *deepfakes*, que elevaram o nível de complexidade das *fake news*, colocando em xeque até mesmo se o que uma pessoa está falando é de fato verdade ou não.

A comunidade científica tem-se empenhado em pesquisar, de forma mais aprofundada sobre as *fake news* e *deepfakes*, nas mais diversas áreas do conhecimento. Ainda mais, tendo em vista que o volume financeiro que vem financiando tal indústria/mercado alcançam cifras significativas, aumentando a presença das *fake news* e *deepfakes* no dia a dia, através das mais diversas mídias de comunicação. Obviamente, esses valores e quem financia esse mercado não são publicizados até por ser uma prática ilegal.

Apesar de serem percebidos resultados satisfatórios, no que tange às soluções contra a criação e propagação de *fake news* e *deepfakes*, a "briga de gato e rato" deixa claro que ainda muito tem a se investigar, assim como buscar formas de minimizar os seus impactos ou mesmo dando mecanismos para que um usuário consiga detectá-las com facilidade. Por isso, os autores deste trabalho procuraram fornecer alguns "insights" de trabalhos futuros que poderão ajudar outros pesquisadores investigarem e encontrarem uma solução definitiva.

Uma abordagem que pode ser utilizada seria uma detecção de *fake news* baseada em plataformas. Poderiam ser criados modelos estatísticos que modelam o comportamento do usuário em busca de maximizar o seu envolvimento com o conteúdo. Esses modelos podem ter em seus algoritmos recursos de autoaprendizagem, utilizando-se de parâmetros que vão se refinando para aumentar a ênfase em busca de informações verdadeiras e que tenham qualidade. Podem ser criados e utilizados algoritmos que apresentam mecanismo de sinalização que consiga ranquear ou indicar se a fonte ou a informação possui credibilidade ou não.

REFERÊNCIAS

Afchar, D., Nozick, V., Yamagishi, J., & Echizen, I. (2018, December). Mesonet: a compact facial video forgery detection network. In *2018 IEEE International Workshop on Information Forensics and Security (WIFS)* (pp. 1-7). IEEE.

Ahmed, H., Traore, I., & Saad, S. (2017, October). Detection of online fake news using n-gram analysis and machine learning techniques. In *International conference on intelligent, secure, and dependable systems in distributed and cloud environments* (pp. 127-138). Springer, Cham.

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of economic perspectives*, *31*(2), 211-36.

Alpaydin, E. (2020). *Introduction to machine learning.* MIT press.

Aphiwongsophon, S., & Chongstitvatana, P. (2018, July). Detecting fake news with machine learning method. In *2018 15th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)* (pp. 528-531). IEEE.

Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, *348*(6239), 1130-1132.

Filho, C. de B., & Calabrez, P. (2017). *Em Busca de Nós Mesmos*. Porto Alegre: Citadel Grupo Editorial.

Basu, S. (1997). The conservatism principle and the asymmetric timeliness of earnings. *Journal of accounting and economics*, *24*(1), 3-37.

Boehm, L. E. (1994). The validity effect: A search for mediating variables. *Personality and Social Psychology Bulletin*, *20*(3), 285-293.

Castells, (2002). *A sociedade em rede*. São Paulo: Paz e Terra.

Castillo, S. I. V., de Oliveira Santos, D., & de Castro, H. C. D. O. (2020). Fake News no contexto da pandemia de COVID-19: considerações a partir da cultura política. *Rizoma*, *8*(1), 165-184.

Colman, A. M. (2015). *A dictionary of psychology*. Oxford quick reference.

Costa, T., 2020. *As 10 redes sociais mais usadas no Brasil [2019]*. [online] Rock Content - BR. Available at: <https://rockcontent.com/br/blog/redes-sociais-mais-usadas-no-brasil/> [Accessed 26 November 2020].

De Carvalho, L. B. (2020). A democracia frustrada: fake news, política e liberdade de expressão nas redes sociais.

De Castro, J. C. L. (2017). A flexibilização da notícia na era dos algoritmos. In *40º Congresso Brasileiro de Ciências da Comunicação*. Curitiba. Retrieved from https://portalintercom.org.br/anais/nacional2017/resumos/R12-2755-1.pdf

De Souza Vieira, L., De Aquino, S. D., & Lins, S. L. B. (2020, September). O que é notícia? Definições que emergem da audiência. In *18º ENCONTRO DA SBPJOR*.

Dias, T., & Oliveira, L. D. (2019). O Monopólio da Verdade na Era das' Fake News'. *Revista Ratio Juris*, *14*(28), 109-126.

Farkas, J., & Schou, J. (2018). Fake news as a floating signifier: Hegemony, antagonism and the politics of falsehood. *Javnost-The Public*, *25*(3), 298-314.

Ferreira, R. R. (2018). Rede de mentiras: a propagação de fake news na pré-campanha presidencial brasileira. *Observatorio (OBS*)*, *12*(5).

Floridi, L. (2018). Artificial intelligence, deepfakes and a future of ectypes. *Philosophy & Technology*, *31*(3), 317-321.

Gabielkov, M., Ramachandran, A., Chaintreau, A., & Legout, A. (2016, June). Social clicks: What and who gets read on Twitter?. In *Proceedings of the 2016 ACM SIGMETRICS international conference on measurement and modeling of computer science* (pp. 179-192).

Ge, C. H. E. N., Jiarong, X. I. E., Guangli, D. A. I., Zheng, P., Xiaqing, H. U., Hongpeng, L. U., ... & Xiaomin, C. H. E. N. (2020). Validity of the use of wrist and forehead temperatures in screening the general population for covid-19: A prospective real-world study. *Iranian Journal of Public Health*.

Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., & Lazer, D. (2019). Fake news on Twitter during the 2016 US presidential election. *Science*, *363*(6425), 374-378.

Güera, D., & Delp, E. J. (2018, November). Deepfake video detection using recurrent neural networks. In *2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)* (pp. 1-6). IEEE.

Guess, A., Nyhan, B., & Reifler, J. (2018). Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 US presidential campaign. *European Research Council*, *9*(3), 4.

Hall, M. (2020, June). Application of the Benford's law to Social bots and Information Operations activities. In *2020 International Conference on Cyber Situational Awareness, Data Analytics and Assessment (CyberSA)* (pp. 1-8). IEEE.

Henry, N., Powell, A., & Flynn, A. (2018). AI can now create fake porn, making revenge porn even more complicated. *The Conversation*, *28*.

Himma-Kadakas, M. (2017). Alternative facts and fake news entering journalistic content production cycle. *Cosmopolitan Civil Societies: An Interdisciplinary Journal*, *9*(2), 25-40.

Huang, D., Zhu, Y., & Mustafaraj, E. (2019, September). How Dependable are" First Impressions" to Distinguish between Real and Fake NewsWebsites?. In *Proceedings of the 30th ACM Conference on Hypertext and Social Media* (pp. 201-210).

Jones, C. (2017). Bill would help California schools teach about 'fake news,' media literacy. Retrieved 30 March 2021, from https://edsource.org/2017/bill-would-help-california-schools-teach-about-fake-newsmedia-literacy/582363

Kahneman, D. (2011). *Rápido e devagar: duas formas de pensar*. Objetiva.

Kong, S. H., Tan, L. M., Gan, K. H., & Samsudin, N. H. (2020, April). Fake News Detection using Deep Learning. In *2020 IEEE 10th Symposium on Computer Applications & Industrial Electronics (ISCAIE)* (pp. 102-107). IEEE.

Korshunov, P., & Marcel, S. (2018). Deepfakes: a new threat to face recognition? assessment and detection. *arXiv preprint arXiv:1812.08685*.

Kuran, T., & Sunstein, C. R. (1998). Availability cascades and risk regulation. *Stan. L. Rev.*, *51*, 683.

Lazer, D. M., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., ... & Zittrain, J. L. (2018). The science of fake news. *Science*, *359*(6380), 1094-1096.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, *521*(7553), 436-444.

Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., ... & Shi, W. (2017). Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4681-4690).

Leite, S., Lourenço, P., & Ferreira, M. (2018). O estado da arte do fenômeno fake news. In *1st Brazil-France-Francophone Belgium Journalism Research Conference*. Retrieved from http://sbpjor.org.br/congresso/index.php/BFFB/sbpjor2018/paper/view/1492/793

Lelo, T. V. (2020). A influência do partidarismo na recepção de fake news e fact-checking em contexto de polarização política. *Observatorio (OBS*)*, *14*(3).

Liv, N., & Greenbaum, D. (2020). Deep fakes and memory malleability: false memories in the service of fake news. *AJOB neuroscience*, *11*(2), 96-104.

Lutz, B., Adam, M. T., Feuerriegel, S., Pröllochs, N., & Neumann, D. (2020). Affective information processing of fake news: Evidence from NeuroIS. In *Information Systems and Neuroscience* (pp. 121-128). Springer, Cham.

Mackenzie, A., & Bhatt, I. (2020). Lies, bullshit and fake news: Some epistemological concerns. *Postdigital Science and Education*, *2*(1), 9-13.

Manohar, S. (2020). Seeing is Deceiving: The Psychology and Neuroscience of Fake Faces.

Messias, J., Schmidt, L., Oliveira, R. A. R., & Souza, F. B. D. (2013). You followed my bot! Transforming robots into influential users in Twitter.

Murziqin, R., Tabrani, Z. A., Idris, S., Bustamam-Ahmad, K., Mendoza, P. J., Huda, M., ... & Qamariah, H. (2020). How to Get Your Research Published and Then Noticed.

Nascimento, C. E. G. (2020). Fake news, mentira organizada e educação: uma reflexão a partir do pensamento de Hannah Arendt. *Revista Docência e Cibercultura*, *4*(1), 243-263.

Oeiras, T., Leite, C., & Castro, R. D. (2018). Os movimentos sociais na internet: a propagação e refutação de Fake News pós-Impeachment. Temática, 14(10). https://doi.org/10.22478/ufpb.1807-8931.2018v14n10.42308

Parikh, S. B., Patil, V., Makawana, R., & Atrey, P. K. (2019, March). Towards impact scoring of fake news. In *2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)* (pp. 529-533). IEEE.

Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you*. Penguin UK.

Passamani, M. S. (2019). *Critérios de classificação de fontes de fake news no brasil e no mundo*. São Paulo: Fundação Getúlio Vargas. Retrieved from https://pesquisa-eaesp.fgv.br/publicacoes/pibic/criterios-de-classificacao-de-fontes-de-fake-news-no-brasil-e-no-mundo

Pennycook, G., & Rand, D. G. (2020). Who falls for fake news? The roles of bullshit receptivity, overclaiming, familiarity, and analytic thinking. *Journal of personality*, *88*(2), 185-200.

Pennycook, G., Cannon, T. D., & Rand, D. G. (2018). Prior exposure increases perceived accuracy of fake news. *Journal of experimental psychology: general*, *147*(12), 1865.

Pessoa, G. P., Botinha, R. M., & Costa, F. D. J. (2018). O Ensino Na Era Da Informação: Um Olhar a Partir Da Neurociência. *Brazilian Journal of Education, Techbology and Society, BRAJETS (11: 4)*, 672-679.

Pierre, L. (1999). Cibercultura. *São Paulo: Editora*, *34*, 264.

Pinto, T. M., & Zanetti, D. (2020). Desinformação e fake news: uma revisão de literatura. *Anais do Seminário Comunicação e Territorialidades*, *1*(6).

Rebs, R. R., & Ernst, A. (2017). Haters e o discurso de ódio: entendendo a violência em sites de redes sociais., *6*(2), 24-44.

Recuero, R., & Gruzd, A. (2019). Cascatas de Fake News Políticas: um estudo de caso no Twitter. *Galáxia (São Paulo)*, (41), 31-47.

Ruchansky, N., Seo, S., & Liu, Y. (2017, November). Csi: A hybrid deep model for fake news detection. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management* (pp. 797-806).

Safieddine, F., & Ibrahim, Y. (Eds.). (2020). *Fake News in an Era of Social Media: Tracking Viral Contagion*. Rowman & Littlefield Publishers.

Schudson, M. (2003). The sociology of news, 2003.

Shae, Z., & Tsai, J. (2019, July). AI blockchain platform for trusting news. In *2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS)* (pp. 1610-1619). IEEE.

Shang, W., Liu, M., Lin, W., & Jia, M. (2018, June). Tracing the source of news based on blockchain. In *2018 IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS)* (pp. 377-381). IEEE.

Shao, C., Ciampaglia, G. L., Varol, O., Flammini, A., & Menczer, F. (2017). The spread of fake news by social bots. *arXiv preprint arXiv:1707.07592*, *96*, 104.

Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data mining perspective. *ACM SIGKDD explorations newsletter*, *19*(1), 22-36.

Silva, G. (2009). O fenômeno noticioso: objeto singular, natureza plural. *Estudos em Jornalismo e Mídia*, *6*(2), 9-15.

Slovic, P., Finucane, M., Peters, E., & MacGregor, D. G. (2002). Rational actors or rational fools: Implications of the affect heuristic for behavioral economics. *The Journal of Socio-Economics*, *31*(4), 329-342.

Souza, C. de F. de O. B. A. de, Quarto, L. C., Souza, A. H. de, Teixeira, F. L. F., Manhães, F. C., & Tiradentes, J. F. V. N. (2020). Um estudo bibliográfico sobre as fake news no âmbito da saúde. In Produção, Comunicação e Representação do Conhecimento e da Informação (pp. 139–145). Atena Editora. https://doi.org/10.22533/at.ed.14620130211

Subramanian, S. (2017). Inside the Macedonian fake-news complex. *Wired magazine*, *15*.

Sunstein, C. R. (2001). *Echo chambers: Bush v. Gore, impeachment, and beyond*. Princeton, NJ: Princeton University Press.

Swan, M. (2015). *Blockchain: Blueprint for a new economy*. " O'Reilly Media, Inc.".

Sydell, L. (2016). We tracked down a fake-news creator in the suburbs. Here's what we learned. *National Public Radio*, *23*.

Tam Cho, W. K., & Gaines, B. J. (2007). Breaking the (Benford) law: Statistical fraud detection in campaign finance. *The american statistician*, *61*(3), 218-223.

Tandoc Jr, E. C., Lim, Z. W., & Ling, R. (2018). Defining "fake news" A typology of scholarly definitions. *Digital journalism*, *6*(2), 137-153.

Tschiatschek, S., Singla, A., Gomez Rodriguez, M., Merchant, A., & Krause, A. (2018, April). Fake news detection in social networks via crowd signals. In *Companion Proceedings of the The Web Conference 2018* (pp. 517-524).

Van der Linden, S., Panagopoulos, C., & Roozenbeek, J. (2020). You are fake news: political bias in perceptions of fake news. *Media, Culture & Society*, *42*(3), 460-470.

Varol, O., Ferrara, E., Menczer, F., & Flammini, A. (2017). Early detection of promoted campaigns on social media. *EPJ Data Science*, *6*, 1-19.

Wardle, C., & Derakhshan, H. (2017). Information disorder: Toward an interdisciplinary framework for research and policy making. *Council of Europe report*, *27*, 1-107.

Wardle, C., & Derakhshan, H. (2018). Thinking about 'information disorder': formats of misinformation, disinformation, and mal-information. *Ireton, Cherilyn; Posetti, Julie. Journalism,'fake news'& disinformation. Paris: Unesco*, 43-54.

Watts, C. (2017). Extremist content and russian disinformation online: Working with tech to find solutions. *Statement prepared for the Senate Judiciary Committee, Subcommittee on Crime and Terrorism.*

Zheng, Z., Xie, S., Dai, H. N., Chen, X., & Wang, H. (2018). Blockchain challenges and opportunities: A survey. *International Journal of Web and Grid Services*, *14*(4), 352-375.

Zhou, X., & Zafarani, R. (2020). A survey of fake news: Fundamental theories, detection methods, and opportunities. *ACM Computing Surveys (CSUR)*, *53*(5), 1-40.