



**UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE ARTES E COMUNICAÇÃO
DEPARTAMENTO DE LETRAS**

BIANCA NATHALLY BEZERRA DE LIMA

**INTELIGÊNCIA ARTIFICIAL E EDUCAÇÃO LINGUÍSTICA: DESAFIOS NA
VERIFICAÇÃO DE *FAKE NEWS* NO CONTEXTO BRASILEIRO**

Recife

2024

BIANCA NATHALLY BEZERRA DE LIMA

**INTELIGÊNCIA ARTIFICIAL E EDUCAÇÃO LINGUÍSTICA: DESAFIOS NA
VERIFICAÇÃO DE *FAKE NEWS* NO CONTEXTO BRASILEIRO**

Trabalho de Conclusão do Curso de Graduação em
Letras: Português – Licenciatura, apresentado como
requisito parcial para a obtenção do grau de Licenciada
em Letras – Português.

Orientadora: Prof^a. Dr^a. Sônia Virgínia Martins Pereira
Coorientador: Prof^o. Dr^o. George Darmiton da Cunha
Cavalcanti

Recife

2024

**Ficha de identificação da obra elaborada pelo autor,
através do programa de geração automática do SIB/UFPE**

Lima, Bianca Nathally Bezerra de.

Inteligência Artificial e Educação Linguística: Desafios na Verificação de Fake News no Contexto Brasileiro / Bianca Nathally Bezerra de Lima. - Recife, 2024.
51 p. : il., tab.

Orientador(a): Sônia Virgínia Martins Pereira

Coorientador(a): George Darmiton da Cunha Cavalcanti

Trabalho de Conclusão de Curso (Graduação) - Universidade Federal de Pernambuco, Centro de Artes e Comunicação, Letras Português - Licenciatura, 2024.

Inclui referências, anexos.

1. Educação. 2. Linguística. 3. Desinformação. 4. Informação. 5. Tecnologia. I. Pereira, Sônia Virgínia Martins. (Orientação). II. Cavalcanti, George Darmiton da Cunha. (Coorientação). IV. Título.

410 CDD (22.ed.)

*À minha mãe Alexsandra e ao meu pai David, que quebraram
um longo ciclo de ignorância e foram os primeiros que
sonharam um futuro resplandecente para mim.*

AGRADECIMENTOS

À minha mãe, Alexsandra, que mesmo diante de críticas e percalços por correr atrás do que era benéfico para mim e meu irmão se manteve resiliente e em quem me espelho. Obrigada por sempre apoiar minhas escolhas na vida.

Ao meu pai, David, que sonhador como sempre foi, montou peça por peça para que tivéssemos nosso primeiro computador e quem me ensinou a amar tudo que amo: músicas, filmes, animações, jogos. O que constitui parte de quem sou.

Ao meu irmão Júlio, que chegou cedo na minha vida, me fazendo companhia e por quem eu daria a vida.

À minhas avós Ivanete Lima e Maria José, e à minha tia Cybelly Souza, mulheres as quais me dão suporte desde que nasci.

À minha amiga Letícia, quem primeiro me acolheu no curso de Letras e muito depois disso. Quem me abrigou e por quem serei sempre grata por termos cruzado caminhos.

Agradeço à minha orientadora Prof. Dra. Sônia Virgínia Martins Pereira por toda sua supervisão durante a jornada de escrita deste trabalho. Quem foi uma verdadeira luz desde o momento que me aceitou como orientanda.

Agradeço também ao Prof. Dr. George Darmiton da Cunha Cavalcanti por ter idealizado o grupo de pesquisa Combate à Desinformação e ao Discurso de Ódio (CoDes) e dado a temática com a qual trabalhei neste projeto e pretendo continuar. Também, por estar sempre aberto a novas ideias e discussões.

Aos meus professores e professoras do Departamento de Letras e do Centro de Educação que me formaram como cidadã e como pessoa. Agradeço especialmente aos coordenadores do Programa Institucional de Bolsas de Iniciação à Docência (PIBID) de Letras Português do ano de 2020: Suzana Leite Cortez, Ricardo Postal e Emanuel Cordeiro da Silva. Também à professora Aline Cunha e aos professores Flaviano Maciel Vieira e Anco Márcio Tenório Vieira. São grandes educadores e marcaram meu caminho na universidade.

A todos os amigos que fiz no projeto Apple Developer Academy UFPE, com quem compartilhei boas conversas durante os almoços no RU.

Por fim, a todos aqueles que, direta ou indiretamente, contribuíram para que eu conseguisse chegar até aqui.

Chego ao final desta etapa tendo mais convicção do que nunca que a maior revolução do pobre é estudar.

"O que significam 70 milhões de anos para seres que só vivem uma milionésima parte desse tempo? Somos como borboletas que pairam no ar durante um dia e pensam que isso é para sempre."

— Carl Sagan, *Cosmos* (1980).

RESUMO

Este estudo investiga os desafios da utilização de inteligência artificial (IA) na detecção de notícias fraudulentas no contexto brasileiro, com foco na educação linguística e no letramento digital. A partir de uma análise qualitativa, são discutidos o papel das plataformas digitais na disseminação de desinformação e as iniciativas de checagem de fatos no Brasil. Ao mobilizar autores como Rojo (2015) e Pires (1998), que destacam a importância dos multiletramentos e da integração das tecnologias ao currículo escolar, o artigo explora a relevância de estratégias educacionais e ferramentas automatizadas no combate à desinformação. Com base nos trabalhos de Wardle e Derakhshan (2023), o estudo também sublinha a importância de ações colaborativas entre instituições educativas, governos, plataformas digitais e a sociedade civil para mitigar os efeitos das notícias fraudulentas. Os resultados indicam que, embora as ferramentas automatizadas de verificação de informações, neste caso os *bots* de checagem de fatos, tenham grande potencial de uso no processo investigativo de notícias fraudulentas, ainda há certas limitações. As ferramentas analisadas mostraram-se mais eficazes quando usadas em contextos específicos. O estudo conclui que a combinação de IA com educação linguística e um esforço colaborativo entre diferentes setores pode ser uma estratégia promissora para combater a desinformação, mas ressalta a importância de continuar aprimorando as tecnologias e métodos de verificação.

Palavras-chave: Inteligência Artificial; Educação Linguística; Letramento Digital; Fake News; Desinformação.

ABSTRACT

This study investigates the challenges of using artificial intelligence (AI) in detecting fraudulent news within the Brazilian context, focusing on linguistic education and digital literacy. Based on a qualitative analysis, the role of digital platforms in disseminating misinformation and fact-checking initiatives in Brazil are discussed. By drawing on authors such as Rojo (2015) and Pires (1998), who emphasize the importance of multiliteracies and the integration of technologies into the school curriculum, the article explores the relevance of educational strategies and automated tools in combating misinformation. Drawing on the work of Wardle and Derakhshan (2023), the study also highlights the importance of collaborative actions between educational institutions, governments, digital platforms, and civil society to mitigate the effects of fraudulent news. The results indicate that although automated fact-checking tools, in this case fact-checking bots, have great potential in the investigative process of fraudulent news, certain limitations remain. The tools analyzed were more effective when used in specific contexts. The study concludes that combining AI with linguistic education and a collaborative effort among different sectors may be a promising strategy to combat misinformation, but emphasizes the importance of continuing to improve verification technologies and methods.

Keywords: Artificial Intelligence; Language Education; Digital Literacy; Fake News; Disinformation.

SUMÁRIO

1 INTRODUÇÃO.....	9
2 BASES CONCEITUAIS PARA A DETECÇÃO DE NOTÍCIAS FRAUDULENTAS: IA E EDUCAÇÃO LINGUÍSTICA.....	15
2.1 Como funcionam os fact-checkers, seu uso e implicações.....	16
2.1.1 Fact-checking automatizado é o remédio para todos os problemas?.....	20
2.1.2 Um problema muito mais profundo.....	24
3 CAMINHOS PARA ANÁLISE DA VERIFICAÇÃO AUTOMATIZADA DE NOTÍCIAS FRAUDULENTAS.....	29
4 À LUZ DO CORPUS: ANÁLISE E DISCUSSÃO.....	31
4.1 Mobilização da educação fronte à desinformação.....	44
5 CONSIDERAÇÕES FINAIS.....	46
REFERÊNCIAS.....	48

1 INTRODUÇÃO

Nos últimos anos, as plataformas digitais emergiram como um meio influente para a disseminação massiva de informações ao redor do mundo, com um impacto particularmente significativo no Brasil (NEWMAN et al., 2019). Esse cenário tem intensificado as preocupações contemporâneas com a desinformação, especialmente quando concentrada no debate político: no Brasil, 85% da população expressa preocupação com a capacidade de distinguir entre o que é factual e o que é enganoso na internet, um dos índices mais altos entre os 38 países pesquisados (NEWMAN et al., 2019). Assim, a "poluição informacional" afeta diversos aspectos do discurso público, incluindo áreas como mudanças climáticas e medicina, na qual, por exemplo, desinformação representa uma ameaça à saúde em todo o mundo (WARDLE; DERAKHSHAN, 2023).

Diante da pandemia de COVID-19, a disseminação de desinformação relacionada à saúde ganhou força, impactando diretamente a aceitação das vacinas (SOUZA FILHO; LAGE, 2021). Notícias fraudulentas sobre os supostos riscos e prejuízos trazidos pelas vacinas circularam amplamente, não só durante a pandemia da COVID-19, mas também em relação a outras vacinas, como a da febre amarela e do sarampo (SOUZA FILHO; LAGE, 2021). No Brasil, apesar da existência do Programa Nacional de Imunizações (PNI), o país sofreu com a reemergência de doenças erradicadas, como o sarampo, em grande parte devido à desinformação disseminada nas redes sociais e ao movimento antivacina, que se fortaleceu por meio da circulação de notícias fraudulentas (SOUZA FILHO; LAGE, 2021). Um exemplo disso foi a disseminação de uma lista intitulada "As 10 razões pelas quais você não deve vacinar seu filho", que reforçou alegações errôneas sobre os malefícios das vacinas, contribuindo para a queda da cobertura vacinal e a reemergência de doenças evitáveis (FERNANDES; MONTOURI, 2020, apud SOUZA FILHO; LAGE, 2021).

Atualmente, com o propósito de combate à desinformação no que tange à área da saúde, o Ministério da Saúde disponibiliza um formulário através da plataforma FalaBr¹ para envio de relatos de recebimento de conteúdo sobre saúde que pode ser falso (Ministério da Saúde, 2023). O formulário solicita o que foi publicado, bem como quem foi o autor da publicação, o link da publicação e, se possível, anexos referentes ao conteúdo em questão. Além disso, o *chatbot* do Ministério da Saúde também é fornecido, através de um número de Whatsapp para esclarecimento de dúvidas.

¹ <https://www.gov.br/acessoainformacao/pt-br/falab>

Diversos esforços e iniciativas de pesquisa têm surgido para mitigar o impacto da disseminação de desinformação online, frequentemente referida como *fake news*. Uma das iniciativas mais amplamente conhecidas é a criação de agências dedicadas à verificação de fatos, ou agências de *fact-checking*, que têm como propósito avaliar a precisão das informações espalhadas nas plataformas digitais (MARQUES et al., 2022, p. 127). Embora os primeiros sites e agências de *fact-checking* tenham surgido no Brasil em 2014, a verificação de fatos na imprensa brasileira ganhou maior visibilidade a partir de 2018, especialmente durante a campanha para a eleição presidencial (SEIBT, 2019, p. 117).

Esse crescimento pode ser ilustrado pela coalizão formada pelo projeto Comprova, que reuniu 24 veículos de mídia para combater a desinformação durante as eleições (SEIBT, 2019, p. 116). Iniciativas como a Agência Lupa, Aos Fatos e Truco também se destacaram, promovendo checagens de declarações públicas, boatos virais e dados eleitorais. Além disso, grandes veículos de comunicação, como o G1 e o jornal Extra, passaram a incluir seções dedicadas exclusivamente ao *fact-checking*, fortalecendo essa prática no ecossistema midiático brasileiro (SEIBT, 2019, p. 28).

Ainda que o uso de *fact-checkers* no Brasil, especialmente aqueles que empregam agentes de inteligência artificial (IA) para analisar *corpus* de diversas fontes, como notícias, redes sociais, blogs etc., tenha crescido — esse crescimento foi particularmente influenciado pelo contexto sociopolítico, especialmente durante as eleições gerais brasileiras de 2014 e 2018 (SEIBT, 2019) —, eles representam apenas uma das ferramentas no combate à desinformação. Quaisquer soluções que se proponham mais robustas e eficientes para o confronto do distúrbio informational requerem ser pautadas em uma abordagem multidisciplinar (WARDLE; DERAKHSHAN, 2023, p. 150).

Wardle e Derakhshan (2023) destacam com firmeza a importância de um "letramento midiático", recomendando, portanto, que não apenas empresas de tecnologia e meios de comunicação estejam envolvidos em suas propostas que visam combater a crescente disseminação de desinformação que permeia a sociedade atual, mas também governos nacionais, ministérios da educação, agências de financiamento e a sociedade civil. Nesse contexto, o cenário brasileiro torna-se particularmente relevante. De acordo com o relatório de notícias digitais de 2019, o Brasil apresenta uma taxa de penetração da internet de 71%, com grande uso de plataformas como Facebook, WhatsApp e YouTube para o consumo de notícias (NEWMAN et al., 2019). Diante desse panorama, é essencial que as estratégias de combate à desinformação considerem a maneira como essas plataformas moldam as relações entre leitura e escrita, exigindo, assim, um repensar crítico dessas práticas.

O manifesto intitulado de "Uma pedagogia dos multiletramentos" (*A Pedagogy of Multiliteracies*), publicado em 1996, como resultado de um colóquio do Grupo de Nova Londres (doravante, GNL), aponta a importância de a escola assumir a responsabilidade pelos novos letramentos que surgem na sociedade contemporânea. Embora os textos em ambientes digitais estejam cada vez mais multissemióticos, multimidiáticos e hipermediáticos, a linguagem escrita continua sendo o principal meio de comunicação. Dessa forma, a escrita tem desempenhado um papel ainda mais central nas interações virtuais, o que torna o estudo dos letramentos digitais um tema de crescente importância (ROJO, 2007).

Com a internet, as barreiras tradicionais entre autor e leitor são dissolvidas, permitindo que os atos de ler e escrever ocorram simultaneamente, uma característica que altera os protocolos de leitura e escrita. A estrutura hipertextual da internet possibilita a criação de novos tipos de textos, onde o leitor também pode ser autor e interagir diretamente com o conteúdo (ROJO, 2007). Essa dinâmica, na qual qualquer pessoa pode atuar como um produtor de conteúdo, é uma das razões pela qual a desinformação se espalha tão rapidamente em plataformas digitais. A facilidade de criar páginas ou grupos nas redes sociais como Facebook e WhatsApp, sem custos iniciais ou requisitos formais, permite que qualquer indivíduo se comporte como um editor de notícias, reforçando o papel da internet como um espaço onde a escrita e a disseminação de informações não estão mais restritas aos especialistas (REIS et al., 2023). Assim também concordam Allcott e Gentzkow (2017) que embora as notícias falsas não sejam um fenômeno novo, as razões para acreditar que sua relevância tem crescido se deve, em parte, à redução das barreiras de entrada na indústria de mídia, com a facilidade de se criar *sites* e monetizar conteúdos por meio de plataformas de publicidade.

Assim, acreditamos que é essencial fornecer às pessoas insumos para: (i) suporte ao entendimento do fenômeno da desinformação considerando diferentes contextos, cenários, e ambientes; (ii) proposição de abordagens automatizadas que possam ser efetivas na contenção e/ou detecção deste tipo de conteúdo, e por fim; (iii) embasamento ao processo de tomada de decisão por parte dos indivíduos até entidades e órgãos de gestão das plataformas digitais e governamentais. De forma geral, esperamos contribuir com esforços focados na minimização dos impactos ocasionados pelo problema em nossa sociedade (REIS et al., 2023).

Além das empresas de notícias tradicionais estarem migrando para plataformas digitais, novos veículos de notícias estão surgindo exclusivamente nesses ambientes (REIS et al., 2023). Esse crescimento de editores de notícias nas plataformas digitais, como exemplificado pelo *Facebook*, reflete uma mudança no meio, mas não necessariamente uma transformação no conteúdo ou na forma como as notícias são produzidas (REIS et al., 2023). De maneira semelhante, a primeira geração das tecnologias de aprendizagem interativas foi uma

transposição do modelo educacional tradicional para o digital, sem grandes mudanças na essência do aprendizado, exceto pela motivação inicial gerada pela novidade (LEMKE, 2010).

Na Finlândia, a integração da checagem de fatos com a educação tem sido uma estratégia fundamental no combate à desinformação. Desde 2014, o governo finlandês reformou seu sistema educacional para priorizar o pensamento crítico e a alfabetização midiática, preparando cidadãos para identificar e lidar com notícias falsas. Esse modelo, apesar de não ser aplicado no Brasil, serve como exemplo de como a educação pode atuar na formação de uma sociedade mais consciente e resistente à desinformação, demonstrando o impacto da checagem de fatos em ambientes educacionais (MACKINTOSH, 2019).

A escola recentemente firmou parceria com a agência finlandesa de checagem de fatos Faktabaari (FactBar) para desenvolver um 'kit de ferramentas' de alfabetização digital destinado a alunos do ensino fundamental ao médio, focado nas eleições da UE. O kit foi apresentado ao grupo de especialistas do bloco em alfabetização midiática e compartilhado entre os estados-membros (MACKINTOSH, 2019)².

Na consciência de que é necessário pensar criticamente acerca da linguagem que empregamos, com o objetivo de compreender de forma efetiva a complexidade dos fenômenos, neste estudo, evitarei o termo *fake news* ou notícia falsa, visto que a expressão *fake news*, nos anos recentes, passou a ser amplamente utilizada, muitas vezes de maneira distorcida, sendo esvaída de seu significado original.

[...] é uma expressão lamentavelmente inadequada para descrever os complexos fenômenos da “poluição” da informação. O termo também começou a ser apropriado por políticos de todo o mundo para descrever organizações de notícias cuja cobertura eles acham desagradável. Dessa forma, está se tornando um mecanismo pelo qual os poderosos podem reprimir, restringir, minar e contornar a liberdade de imprensa. (WARDLE; DERAKHASHAN, 2023, p. 5).

Eugenio Bucci, professor da Universidade de São Paulo, sugere uma tradução mais precisa para o termo *fake news*, alinhada com a proposta de Carlos Eduardo Lins da Silva, que utiliza a expressão "notícia fraudulenta" para descrever o fenômeno (SEIBT, 2019, p. 99).

Não se resume à publicação de informações incorretas ou informações não factuais. Nós estamos diante de uma usina de produção de notícias fraudulentas, que são forjadas com aparência de ser jornalística confiável, mas não são, com o propósito de fraudar os processos decisórios das democracias. Isso quer dizer que a expressão *fake news* designa uma notícia fabricada com má intenção, que se vale do aspecto de uma notícia jornalística com o propósito de enganar o público. É muito diferente, portanto, de um erro jornalístico, coisas que acontecem todo dia. Uma boa redação jornalística quando comete um erro, ela procura se corrigir (BUCCI, 2018).

² The school recently partnered with Finnish fact-checking agency Faktabaari (FactBar) to develop a digital literacy “toolkit” for elementary to high school students learning about the EU elections. It was presented to the bloc’s expert group on media literacy and has been shared among member states.

Acerca do termo "inteligência artificial"³, Buzato (2023) argumenta que o termo é muitas vezes utilizado de forma antropocêntrica, atribuindo características humanas a artefatos que, na verdade, operam de maneira técnica e não-consciente. Ele sugere que essa personalização da IA, como a que ocorre em modelos de linguagem natural, não é produtiva do ponto de vista educacional. Em vez disso, o autor propõe o modelo de "assemblagem cognitiva", no qual as interações com a IA não devem ser confundidas com consciências humanas, mas vistas como processos técnicos que influenciam modos de atenção e reconhecimento. Apesar disso, utilizarei o termo inteligência artificial (IA) como um termo guarda-chuva para designar as técnicas de automatização e aprendizagem utilizadas na detecção de notícias fraudulentas.

Diante do exposto, surge a questão central: como o letramento digital e o uso de ferramentas de IA podem contribuir de maneira eficaz para a detecção de notícias fraudulentas no contexto brasileiro? O propósito do presente trabalho é investigar os desafios e potencialidades do uso de *fact-checkers* na detecção de notícias fraudulentas, no contexto brasileiro, com foco no impacto da educação linguística e do letramento digital para esse fim. Para tanto, pretende-se (1) verificar o papel dos *fact-checkers* e sua eficiência, (2) discutir a importância de um letramento digital nas escolas brasileiras e (3) examinar o uso de abordagens automatizadas (tabela 1) na contenção da desinformação que possam ser utilizadas por estudantes. Trata-se, portanto, de uma pesquisa qualitativa, com natureza exploratória e descritiva, que busca examinar conceitos e analisar dados relevantes sobre o tema do estudo.

³ No que diz respeito à perspectiva computacional, a definição de Inteligência Artificial (IA) apresentada no prefácio de Russell e Norvig (2010) destaca a IA como um campo abrangente que envolve a aplicação de lógica, probabilidade e matemática contínua para abordar processos de percepção, raciocínio, aprendizado e ação. O objetivo do estudo da IA é a criação de agentes inteligentes capazes de perceber o ambiente e agir de forma racional para atingir metas. Além disso, os autores ressaltam que a IA tem evoluído ao longo do tempo, incorporando avanços teóricos e práticos, como reconhecimento de fala, aprendizado de máquina e visão computacional, demonstrando que a área não se limita a subcampos isolados, mas busca integrar conhecimentos de forma abrangente e moderna (Russell & Norvig, 2010).

Tabela 1 - Abordagens automatizadas utilizadas para análise

Ferramenta	Propósito	Metodologia de verificação	Fontes de dados e datasets	Características adicionais
Bot do "Aos Fatos", Fátima 3.0⁴	Verificar informações em tempo real, principalmente em redes sociais como o WhatsApp e o Telegram. Também é possível utilizar o <i>bot</i> no navegador.	Utiliza inteligência artificial e aprendizado de máquina para identificar e verificar informações compartilhadas pelos usuários, com base em banco de dados de notícias verificadas previamente.	Utiliza a base de dados do "Aos Fatos", que inclui notícias verificadas manualmente e contribuições de usuários; alimentado por informações de eventos importantes e dados de desinformação identificados pela equipe.	Realiza verificações em tempo real e permite que os usuários encaminhem mensagens para análise, sendo eficiente em ambientes de comunicação instantânea.
Detector de notícias fraudulentas FakeCheck⁵	Auxiliar na verificação de notícias fraudulentas em português, identificando padrões de desinformação em textos e manchetes.	Utiliza algoritmos de processamento de linguagem natural (PLN) e aprendizado de máquina para analisar e classificar notícias como "verdadeiras" ou "falsas" através de duas classificações: palavras do texto ou classe gramatical.	Utilizou datasets como o "Fake.Br Corpus" e outros conjuntos de dados de notícias fraudulentas e "verdadeiras" em português para treinar o modelo.	Tem foco em verificar textos e é particularmente útil para análise de notícias textuais, identificando padrões linguísticos comuns em notícias fraudulentas.

Fonte: A autora (2024)

Este trabalho está organizado em cinco seções. Após esta introdução, a segunda seção apresenta a fundamentação teórica, a qual discute abordagens conceituais, como o funcionamento dos *fact-checkers* e suas implicações, além da importância do letramento midiático. A terceira seção detalha a metodologia utilizada e o *corpus* de análise. Na quarta, ocorre a análise do *corpus* com base nos conceitos apresentados. Por fim, as considerações finais retomam os resultados obtidos e as reflexões sobre o problema e os objetivos da pesquisa.

⁴ <https://fatimagpt.aosfatos.org/>

⁵ <https://nilc-fakenews.herokuapp.com/>

2 BASES CONCEITUAIS PARA A DETECÇÃO DE NOTÍCIAS FRAUDULENTAS: IA E EDUCAÇÃO LINGUÍSTICA

A discussão sobre a fragmentação do conhecimento e a necessidade de integração entre disciplinas já era uma preocupação no final do século XX. Pires (1998) trata sobre como essa fragmentação, herdada do modelo industrial de organização do trabalho, reflete-se diretamente no ensino, que permanece compartmentado e desarticulado. Embora haja uma necessidade crescente de integração, os currículos escolares no Brasil ainda operam de forma estanque, o que impede uma formação mais crítica e capaz de responder às demandas complexas da sociedade contemporânea. Pires argumenta que essa situação limita o potencial transformador do ensino e propõe que a superação dessa fragmentação exige não apenas a junção superficial de disciplinas, mas uma articulação mais profunda entre os caminhos epistemológicos e metodológicos, promovendo uma formação mais integral (PIRES, 1998).

O currículo escolar, ao ser elaborado, deve refletir criticamente as demandas sociais, proporcionando aos alunos experiências que os preparem para participar de forma responsável e crítica nas práticas da *web* (ROJO, 2015, P. 135). Nesse contexto, é fundamental considerar o conceito de multiletramentos, que, segundo o Grupo de Nova Londres, envolve duas principais multiplicidades: a diversidade cultural, representada pelo multiculturalismo, e a variedade de linguagens e mídias, ou multissemiose (ROJO, 2015, p. 135). A partir desse gancho, Rojo (2015) aborda a forma como a escola ainda favorece predominantemente a cultura tradicional, deixando de considerar os multi e novos letramentos e as práticas que circulam em ambientes digitais e de cultura de massa.

[...] a escola [...] deve propiciar experiências significativas com produções de diferentes culturas e com práticas, procedimentos e gêneros que circulam em ambientes digitais: refletir sobre participações, avaliar a sustentação das opiniões, a pertinência e adequação de comentários, a imagem que se passa, a confiabilidade das fontes, apurar os critérios de curadoria e de seleção de textos/produções, refinar os processos de produção e recepção de textos multissemióticos (ROJO, 2015, p. 135).

Segundo Cavalcante e Souza (2016), em tempos de cibercultura a leitura e suas relações com o mundo midiático assumem novas vertentes, gestos e posturas. O surgimento de novos gêneros textuais e suas formas de produção, reprodução e apropriação pelo leitor tornam-se evidentes, especialmente com a presença das TDIC, Tecnologias Digitais de Informação e Comunicação (CAVALCANTE; SOUZA, 2016). Lemke (2010) discute que, ao especificar os recursos semióticos e tecnológicos que caracterizam cada letramento, é possível identificar tantos letramentos quanto os gêneros multimidiáticos. Além disso, ele observa que essas categorias podem ser ampliadas ao considerar competências tanto na produção quanto no uso das tecnologias. No passado, com o uso de materiais simples como caneta, papel e livros, ou

até mesmo máquinas de datilografia, essas distinções eram mais claras. Entretanto, hoje, ao se lidar com hipertextos, tanto para lê-los quanto escrevê-los, conforme Bolter (1998), são necessárias diferentes tipos de tecnologias, além de novas habilidades de autoria e interpretação para usá-las.

Aquilo que parece ser o mesmo texto ou gênero multimidiático não é funcionalmente o mesmo quando no papel ou na tela, segue diferentes convenções de significado e requer diferentes habilidades para que seu uso seja bem sucedido, quando funciona em diferentes redes sociais para diferentes objetivos, como parte de diferentes atividades humanas. Um letramento é sempre um letramento em algum gênero e deve ser definido com respeito aos sistemas sígnicos empregados, às tecnologias materiais usadas e aos contextos sociais de produção, circulação e uso de um gênero particular. Podemos ser letrados em um gênero de relato de pesquisa científica ou em um gênero de apresentação de negócios. Em cada caso as habilidades de letramento específicas e as comunidades de comunicação relevantes são muito diferentes (LEMKE, 2010, p. 457).

De acordo com Lemke (2010), a publicação na internet exige uma série de conhecimentos e habilidades que vão além da simples escrita e digitação. Para que um trabalho seja adequadamente formatado e disponibilizado online, é necessário operar e compreender aspectos do *software* e *hardware*, bem como lidar com a manutenção da rede e o desenvolvimento de protocolos. Embora o computador funcione como um processador de informação universal e simplifique o processo, há a necessidade de um conhecimento mais amplo e diversificado, tanto em termos materiais quanto semióticos, para projetar, manter e utilizar efetivamente esse sistema de práticas.

Garcia (2020, p. 21-22) argumenta que avanços significativos foram realizados recentemente na pesquisa técnica e multidisciplinar, mas ainda é necessário maior investimento. Além disso, o apoio de governos, centros de pesquisa e líderes empresariais é fundamental para promover o progresso da IA ética, possibilitando a análise externa dos dados utilizados. A implementação consciente e menos tendenciosa de inteligência artificial demanda uma abordagem multidisciplinar, envolvendo especialistas em ética, ciências sociais e outras áreas relacionadas às aplicações específicas da tecnologia (GARCIA, 2020, p. 22).

2.1 Como funcionam os *fact-checkers*, seu uso e implicações

A verificação de fatos — ou *fact-checking* — é a tarefa de avaliar se as afirmações feitas em linguagem escrita ou falada são factuais (GUO et al., 2022, p. 178)⁶. Farhangian et al. (2024) apontam que a detecção automática de notícias fraudulentas tornou-se uma área emergente de pesquisa dentro do aprendizado de máquina e da inteligência artificial. Devido ao fato de que a detecção desse tipo de notícia desinformativa com base em características linguísticas, como o conteúdo textual de notícias ou postagens em redes sociais, é uma das

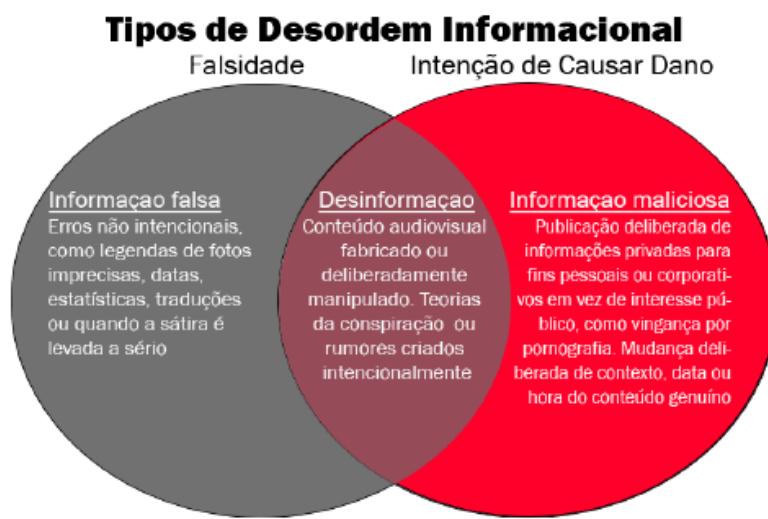
⁶ Fact-checking is the task of assessing whether claims made in written or spoken language are true.

abordagens mais atrativas, a maioria dos estudos comparativos foca nesse aspecto (FARHANGIAN et al., 2024).

Ao utilizar o termo "factual", entende-se não apenas como uma simples referência ao que é verdadeiro, mas como parte de uma discussão mais ampla sobre os regimes de verdade. Conforme argumenta Bentes (2019), em sua palestra "Fake News, narrativas e memética", apresentada no evento XIV Abralin em Cena: Fake News e Linguagem, o conceito de verdade não é algo estático ou absoluto, mas sim construído e disputado em diferentes contextos sociais e culturais. Bentes destaca que as verdades são moldadas por crenças, consensos científicos e narrativas culturais, e que a questão da verdade é, na realidade, uma questão ética e política. Mais adiante, no tópico 2.1.2, será aprofundada a discussão sobre como a verdade se configura como um conceito em constante disputa e modulação.

De início, se faz necessário diferenciar em escopo a verificação de fatos e a detecção de notícias fraudulentas. Segundo Guo et al. (2022), enquanto a verificação de fatos se concentra em avaliar a veracidade de afirmações específicas, a detecção de notícias desinformativas tem um foco mais amplo, abordando artigos inteiros, incluindo aspectos que não se limitam à veracidade, como a detecção de sátira. Além disso, a detecção de notícias falsas pode considerar fatores como o público-alvo e as intenções do autor.

Figura 1 - Intersecção conceitual e o conceito de desinformação.



Fonte: Wardle e Derakhshan (2023)

Apesar de outros estudos categorizarem diversificadas tipologias, como a de Molina et al. (2019), na qual os conteúdos foram divididos em oito tipos, neste trabalho considero a de Wardle e Derakhshan (2023). Os autores apresentam três tipos de desordem informacional

que são fundamentais para compreender como a desinformação se manifesta e se propaga na sociedade contemporânea. Primeiro, eles destacam a informação falsa (*misinformation*), que ocorre quando informações incorretas ou enganosas são compartilhadas sem a intenção de causar danos. Em segundo lugar, está a desinformação (*disinformation*), caracterizada por informações falsas disseminadas deliberadamente com a intenção de causar danos. Por fim, há a informação maliciosa (*malinformation*), caracterizada pela divulgação de informações factuais com a intenção de prejudicar, frequentemente expondo dados privados ao público. A figura 1 examina como a informação falsa, desinformação e informação maliciosa se interseccionam em relação aos conceitos de falsidade e intenção de causar dano (Wardle e Derakhshan, 2023, p. 46).

Certos tipos de discurso, como o de ódio e o assédio, são classificados como informação maliciosa, pois têm como alvo indivíduos com base em suas características pessoais ou afiliações. Embora essas informações possam se basear em fatos reais, elas são usadas de forma estratégica para causar danos (Figura 1). Os autores ressaltam que a distinção entre esses tipos de desordem informacional é crucial para entender a complexidade da propagação de informações no cenário atual e para desenvolver estratégias eficazes de combate à desinformação. Dada a perspectiva acerca da disseminação de informações fraudulentas deste trabalho, o foco da abordagem tende para a informação falsa e desinformação.

Isto posto, antes de discutir mais a fundo os benefícios e limitações da inteligência artificial, é necessário compreender seu funcionamento básico (GARCIA, 2020, p. 15). A IA é uma área da computação voltada ao desenvolvimento de algoritmos e sistemas capazes de realizar tarefas que exigem habilidades associadas à inteligência humana e também abrange a comunicação em linguagem natural, como os assistentes pessoais, e o reconhecimento de imagens, como nos sistemas de carros autônomos (GARCIA, 2020, p. 15). Além disso, as técnicas de IA permitem que máquinas planejem, raciocinem e resolvam problemas complexos de forma autônoma, como exemplificado por sistemas de diagnóstico médico (GARCIA, 2020, p. 15).

Garcia (2020) define aprendizado de máquina, ou como é comumente referido *machine learning*, como uma subárea da IA, a qual envolve a capacidade das máquinas de aprender com exemplos e dados. Nesse modelo, em vez de programar manualmente cada etapa de um processo, o algoritmo é treinado para identificar padrões e soluções com base em grandes conjuntos de dados. Isso permite que a máquina execute tarefas novas e complexas sem a necessidade de instruções específicas para cada caso. Durante o treinamento, os algoritmos geram modelos matemáticos que mapeiam entradas e saídas, ajustando os coeficientes das

equações para generalizar as soluções. Um modelo ideal é aquele que não apenas “decora” os dados de treinamento, mas que consegue extrair padrões que possam ser aplicados a novos problemas (GARCIA, 2020, p. 16).

Sem dados e outros incrementos [...] a IA não teria o sucesso que tem. Os dados são fornecidos pelos usuários nas variadas redes da internet, não apenas as sociais. A internet agora gera e distribui novos dados em quantidades exorbitantes. [...] sendo a maior parte de dados não estruturados⁷, como vídeos e áudios. Todos esses dados estão relacionados com atividades humanas, sentimentos, experiências e relações do dia a dia. [...] Sem a IA, embora não possamos considerá-los como dados neutros, esses dados não passariam de dados mortos, ou melhor, inconsequentes (SANTAELLA, 2023, p. 15).

Segundo Garcia (2020), apesar dos avanços proporcionados pela inteligência artificial, um dos grandes desafios enfrentados é o viés nos dados utilizados para treinar os algoritmos. Os dados não são neutros, pois refletem escolhas e decisões humanas que, muitas vezes, estão carregadas de preconceitos. Se os dados que alimentam os sistemas contêm vieses de raça, gênero ou outras características, as máquinas não apenas aprenderão esses padrões, mas também os perpetuarão durante o seu processo de aprendizado (GARCIA, 2020, p. 17). Em áreas mais sensíveis, como decisões judiciais e de saúde, esses vieses podem ter impactos ainda mais significativos. No caso do sistema COMPAS⁸, utilizado nos Estados Unidos da América para decisões de liberdade condicional, o viés nos dados reforçou estereótipos raciais, resultando em maior severidade para réus negros em comparação aos brancos. Esse exemplo ilustra como a falta de cuidado no uso de dados enviesados pode perpetuar desigualdades, especialmente em sistemas automatizados que recebem um status de imparcialidade apenas pela sua natureza matemática (GARCIA, 2020, p. 18). Cathy O'Neil (2016) destaca a problemática da discriminação reforçada por *big data* e algoritmos matemáticos, ao afirmar que:

[...] se um estudante pobre não consegue um empréstimo porque um modelo de crédito o considera muito arriscado (com base no seu código postal), ele é então privado do tipo de educação que poderia tirá-lo da pobreza, e um ciclo vicioso se inicia. Os modelos estão sustentando os privilegiados e punindo os desfavorecidos, criando um 'coquetel tóxico para a democracia' (O'NEIL, 2016)⁹.

Através de Guo et al. (2022) pode-se compreender que o uso de técnicas de Processamento de Linguagem Natural (PLN) e aprendizado de máquina tem se tornado fundamental para a automação do processo de checagem de fatos, devido à rapidez com que a informação e a desinformação se espalham no ecossistema de mídia moderno. Essas técnicas

⁷ Os dados estruturados são mais facilmente manipuláveis por sistemas computacionais devido à existência de linguagens formais, como SQL, que permitem consultas e manipulações precisas. Já os dados não estruturados exigem mecanismos computacionais distintos dos métodos tradicionais para serem processados, sendo necessário impor alguma estrutura para que métodos tradicionais de mineração de dados possam ser aplicados (SCARINCI; PALAZZO, 2001, p. 15).

⁸ Software norte-americano que tem a finalidade de verificação da possibilidade de reincidência criminal

⁹ [...] if a poor student can't get a loan because a lending model deems him too risky (by virtue of his zip code), he's then cut off from the kind of education that could pull him out of poverty, and a vicious spiral ensues. Models are propping up the lucky and punishing the downtrodden, creating a 'toxic cocktail for democracy'.

possibilitam a previsão da veracidade das afirmações de maneira mais eficiente, lidando com grandes volumes de dados e identificando padrões relacionados à desinformação de forma mais precisa.

O processo de automação da verificação de fatos tem sido amplamente discutido no contexto do jornalismo computacional e tem recebido atenção significativa na comunidade de inteligência artificial (GUO et al., 2022, p. 178). Pesquisadores como Vlachos e Riedel (2014) propuseram uma abordagem estruturada, composta por etapas como a identificação de afirmações a serem verificadas, a busca de evidências adequadas e a formulação de veredictos, que podem ser modeladas como tarefas de Processamento de Linguagem Natural (PLN)¹⁰. Essa abordagem incentivou o desenvolvimento de *pipelines*¹¹ automatizadas baseadas em tarefas já exploradas pela comunidade de PLN, com avanços possibilitados por *datasets*¹² criados para pesquisa.

2.1.1 *Fact-checking* automatizado é o remédio para todos os problemas?

Considerando o cenário do estudo sobre a detecção automática de notícias fraudulentas em português, Monteiro et al. (2018) desenvolveram uma pesquisa que visava criar o primeiro *corpus* de referência de notícias "verdadeiras" e "falsas" — sendo a terminologia utilizada pelos autores — na língua portuguesa, conhecido como "Fake.Br Corpus". Este trabalho em questão aborda a crescente problemática das notícias fraudulentas e para contribuir com a identificação e combate a essas informações enganosas, os autores seguiram várias etapas em sua metodologia. O processo começou com a coleta de dados, que envolveu a busca manual de notícias fraudulentas a partir de quatro diferentes *websites*: Diário do Brasil (3,338 notícias), A Folha do Brasil (190 notícias), The Jornal Brasil (65 notícias) e Top Five TV (7 notícias), e a coleta semiautomática de notícias factuais. Em seguida, ocorreu o alinhamento de notícias, assegurando que cada notícia fraudulenta tivesse uma correspondente factual sobre o mesmo tema, possibilitando melhores comparações. A etapa subsequente se caracterizou pela análise de características linguísticas, onde os autores investigaram diversos aspectos:

(i) pausalidade, que verifica a frequência de pausas em um texto, calculada como o número de sinais de pontuação em relação ao número de frases; (ii) emotividade, que é uma indicação da expressividade da linguagem em uma mensagem, calculada como a soma do número de adjetivos e advérbios em relação à soma de

¹⁰ De forma breve, integra áreas como ciência da computação, inteligência artificial e linguística, voltando-se para a produção e compreensão automatizada da linguagem natural.

¹¹ Consiste em um conjunto de procedimentos que devem ser seguidos para implementar a aplicação em produção.

¹² Trata-se de uma coleção de dados estruturados, reunidos e organizados para fins de análise ou processamento.

substantivos e verbos; (iii) incerteza, medida pelo número de verbos modais e ocorrências de vozes passivas; e (iv) não-immediatismo, medida pelo número de pronomes de 1^a e 2^a pessoa (MONTEIRO et al., 2018, p. 5-6)¹³.

Com o *corpus* estabelecido, os autores passaram para o treinamento e avaliação de modelos de aprendizado de máquina, aplicando diferentes conjuntos de características linguísticas para detectar automaticamente as notícias fraudulentas. A última parte consistiu na avaliação de resultados, utilizando determinadas métricas para validar a eficácia dos modelos. Monteiro et al. (2018) dividiram as notícias coletadas entre "*fake news*" e "*true news*" dentro de seu *corpus*. Ou seja, meias verdades foram excluídas da classificação.

Embora o processo de treinamento e avaliação dos modelos tenha sido conduzido com base na classificação entre *fake* e *true news*, como proposto por Monteiro et al. (2018), é importante reconhecer que essa abordagem binária pode não capturar toda a complexidade da desinformação. Na prática, notícias fraudulentas nem sempre têm conteúdo totalmente destituído de factualidades ou totalmente inverossímeis; muitas vezes, elas incorporam elementos de lastro factual misturados com desinformação, imprecisões ou até mesmo erros não intencionais, tornando a classificação binária mais desafiadora.

Por exemplo, uma notícia que contenha um erro jornalístico involuntário, a qual se encaixaria na classificação de informação falsa descrita por Wardle e Derakhshan (2023), seria definida como "*fake news*" por um classificador automático da mesma forma que uma notícia totalmente fabricada. Wardle e Derakhshan (2023) sugerem que, em um cenário real, a informação enganosa nem sempre é totalmente inverossímil. Eles apontam que mensagens podem variar em precisão, desde manchetes que distorcem o conteúdo até informações 100% fabricadas. A informação maliciosa, por exemplo, pode usar dados verdadeiros para causar dano ou manipular a percepção pública (WARDLE; DERAKHSHAN, 2023, p. 57). Além disso, "as tecnologias para análise automática de textos são significativamente diferentes das tecnologias de análise de imagens estáticas e em movimento" (WARDLE; DERAKHSHAN, 2023, p. 57). A desinformação pode adotar diversos formatos (texto, imagem, vídeo) e, muitas vezes, apresenta-se como uma fonte oficial ou utiliza elementos verdadeiros para ganhar credibilidade, ressaltando a complexidade de se identificar o que é factual ou não (WARDLE; DERAKHSHAN, 2023, p. 57).

[...] se uma mensagem é parcial ou totalmente falsa, mas nenhum dano é pretendido por seu produtor, ela não se enquadra na definição de desinformação. Por essa razão, é importante diferenciar entre informação falsa (falsa, mas sem intenção de causar dano) e informação maliciosa (verdadeira, mas com intenção de causar dano) (WARDLE; DERAKHSHAN, 2023, p. 69).

¹³ (i) pausality, which checks the frequency of pauses in a text, computed as the number of punctuation signals over the number of sentences; (ii) emotiveness, which is an indication of language expressiveness in a message [23], computed as the sum of the number of adjectives and adverbs over the sum of nouns and verbs; (iii) uncertainty, measured by the number of modal verbs and occurrences of passive voices; and (iv) nonimmediacy, measured by the number of 1st and 2nd pronouns.

Levando esses aspectos em consideração e destacando que a questão abordada acerca da classificação não é exclusiva do *corpus* Fake.Br, mas utilizada aqui como exemplo, embora a classificação de *datasets* ofereça uma base consistente para a detecção automatizada de notícias fraudulentas, ela pode simplificar a complexidade da desinformação em diversos contextos, deixando de capturar nuances importantes que são essenciais para a compreensão completa do fenômeno.

Cada agência de checagem de fatos utiliza seus próprios rótulos para classificar as alegações. Os rótulos podem ser, por exemplo: falso, verdadeiro, impossível de provar, distorcido, exagerado, controverso, sem contexto, impreciso, entre outros. No ClaimReview, a classificação da alegação deve ser realizada em uma escala linear, onde 1 representa falso e o número mais alto da escala, definido pelo valor em BestRating, representa verdadeiro, com o intervalo indicando meias-verdades (MORENO; BRESSAN, 2019, p. 526)¹⁴.

Em razão destes e de outros fatores já mencionados, há a necessidade de uma educação midiática, como já mencionado anteriormente. Wardle e Derakhshan (2023) sugerem que um currículo eficaz deve incluir diversos elementos como habilidades tradicionais de educação midiática, técnicas investigativas para verificação de mídia social, entendimento do poder dos algoritmos na modelagem das informações que recebemos, além das possibilidades e implicações éticas da inteligência artificial. Eles também destacam a importância de desenvolver ceticismo emocional para evitar a falta de criticidade diante de conteúdos que provocam respostas emocionais, bem como o aprimoramento de habilidades em cálculo estatístico.

Em consonância com essa necessidade de uma educação midiática mais abrangente, Rojo (2017), destaca que a educação contemporânea enfrenta o desafio de adaptar-se à natureza multimidiática da sociedade atual. A expressão de ideias e o acesso ao conhecimento não se limitam mais ao texto escrito; na realidade, a vida e a comunicação tornaram-se predominantemente multimídias (ROJO, 2017, p. 4). Portanto, é urgente incorporar os multiletramentos e novos letramentos ao currículo escolar, visto que a maioria dos gêneros discursivos utilizados hoje em dia inclui não apenas textos escritos, mas também imagens, diagramas, tabelas, vídeos, sons e diversas outras formas de linguagem (ROJO, 2017, p. 4). Deste modo, as habilidades tradicionais de leitura e escrita já não são suficientes para atender às demandas da vida contemporânea, ressaltando a necessidade de que os currículos escolares sejam atualizados para refletir essa realidade (ROJO, 2017, p. 4).

Os leitores que crescem em uma cultura digital e colaborativa apreendem novos gestos e habilidades, se apropriando e, ao mesmo tempo, criando elementos que condizem com os diferentes discursos travados no espaço virtual, por exemplo, utilizando-se de ícones (emoticons e sinais gráficos) ou de uma escrita abreviada

¹⁴ Each fact-check agency uses its own labels to classify claims. The Labels can be for example: false, true, impossible to prove, distorted, exaggerated, controversial, without context, inaccurate, among others. In the claimReview the classification of the claim should be performed on a linear scale, where 1 is false and the higher number at the scale, set by the value in BestRating is true, with the interval representing half trues.

para se comunicar. É fato que o leitor dispõe de mais dispositivos para ler e modos de fazê-lo do que em épocas anteriores, ampliando, assim, as possibilidades de leitura e de apropriação do que é lido. Antes de ser visto como um ponto negativo, esses espaços possibilitam, ainda mais, a democratização do acesso à informação (CAVALCANTI; SOUZA, 2016).

Essa perspectiva acerca de multimídias se alinha ao que o diretor da Fact-Checking Network, Alexios Mantzarlis (WARDLE; DERAKHSHAN, 2023), aponta sobre a necessidade de adaptar formatos para desmascarar informações não factuais e engajar um público mais amplo. Mantzarlis assinala que é essencial encontrar formatos mais envolventes para pessoas que evitam artigos longos e *hiperlinks*, e menciona o sucesso de plataformas como *Snapchat* [à época] e *bots*, embora ainda existam lacunas, como a ausência de um *podcast* inovador ou conflitos com a televisão.

Nascimento et al. (2013) ao abordar o discurso de ódio citam como as plataformas de mídia social, como *Facebook*, *Twitter* e *YouTube* afirmaram ter a intenção de resolver esse problema por meio de políticas voltadas ao comportamento de ódio; no entanto, grande parte da moderação de conteúdo ainda depende de revisão manual de documentos questionáveis. Esse processo é considerado trabalhoso, demorado, caro e não escalável, especialmente devido à velocidade com que as mensagens são transmitidas (NASCIMENTO et al., 2023).

Há dois anos, nossa produção diária de dados [na internet] estava estimada em torno de 2,5 quintilhões (um número com dezoito zeros) de bytes [...]. [...] por dia, o Google realizava 3.607.080 pesquisas, os usuários do Twitter enviavam 456 mil tuítes, os usuários do YouTube assistiam a 4.146.600 vídeos, as caixas de entrada recebiam 103.447, 529 e-mail spam, o canal do tempo recebia 18.055.555.56 solicitações de clima. (MUELLER e MASSARON, 2020, p. 24, 26, 35). Isso é big data (SANTAELLA, 2023, p. 15).

Seibt (2019) descreve como o *fact-checking*, ou jornalismo de verificação, surgiu como uma resposta às crescentes mudanças na dinâmica da comunicação pública e à proliferação da desinformação e como essa prática resgata valores centrais, como a objetividade, pelo menos em teoria, e a disciplina da verificação. Ao trazer maior transparência e rigor à apuração de fatos, o *fact-checking* desempenha um papel fundamental no combate à desinformação (SEIBT, 2019, p. 18-20).

Além disso, o *fact-checking* tem o potencial de influenciar positivamente o comportamento dos agentes políticos ao impor custos reputacionais para a disseminação de desinformação (SEIBT, 2019). A questão mais relevante, como destacado, não é apenas se as informações serão desmentidas, mas sim "se declarações mais enganosas ou imprecisas seriam feitas na ausência de verificadores de fatos" (NYHAN; REIFLER, 2014, p. 4, citado em SEIBT, 2019). Seibt relata que Nyhan e Reifler (2014) conduziram uma pesquisa em nove estados americanos antes das eleições de 2012 para investigar o impacto do *fact-checking* no Congresso, enviando alertas sobre a presença do PolitiFact a um grupo de legisladores. A pesquisa revelou que o grupo que recebeu o alerta sobre verificação apresentou uma menor

prevalência de avaliações negativas em comparação com os grupos controle e placebo, sugerindo que o *fact-checking* "[...] pode ter como efeito colateral a responsabilização de agentes políticos com relação às suas declarações, aumentando o custo político da desinformação" (SEIBT, 2019).

2.1.2 Um problema muito mais profundo

À luz destas circunstâncias, dado o contexto em que este trabalho está inserido, no que diz respeito à multidisciplinaridade e à abordagem de notícias fraudulentas, é fundamental tratar do que é verdade. Ivana Bentes (2019), em sua fala, intitulada "Fake News, narrativas e memética", apresentada no evento XIV Abralin em Cena: Fake News e Linguagem, argumenta sobre como ao discutir acerca de *fake news* é necessário entender outras crises. Bentes liga diretamente o fenômeno das notícias fraudulentas à uma crise do próprio conceito de verdade, inserindo na discussão Nietzsche, em Crepúsculo dos Ídolos (2006, p. 32), que afirma: "O 'mundo verdadeiro' – uma ideia que para nada mais serve, não mais obriga a nada – ideia tornada inútil, logo refutada: vamos eliminá-la!", ressaltando que o conceito de verdade nunca foi tão eficiente para disputar valores. Além disso, a palestrante também trata na conceituação acerca dos regimes de verdade de Foucault (1979):

Cada sociedade tem seu regime de verdade, sua "política geral" de verdade, isto é, os tipos de discurso que aceita e faz funcionar como verdadeiros..., os meios pelo qual cada um deles é sancionado, as técnicas e procedimentos valorizados na aquisição da verdade; o *status* daqueles que estão encarregados de dizer o que conta como verdadeiro (FOUCAULT, 1979).

Bentes (2019) defende que a verdade não deve ser entendida como um conceito central, pois, em meio a crenças, consensos científicos e a compreensão de que as verdades são construídas, disputadas e conceituais, a questão essencial deixa de ser se algo é verdadeiro ou falso, tornando-se, antes, uma questão ética. Ela exemplifica que as narrativas religiosas, as mitologias nacionais, as convenções econômicas e teorias científicas representam verdades temporárias, sujeitas a serem refutadas a qualquer momento (BENTES, 2019). As agências de checagem trazem uma novidade no que diz respeito às formas de combater notícias fraudulentas no que tange à classificar notícias de forma mais abrangente do que apenas em verdadeiro ou falso (BENTES, 2019).

A Agência Lupa, por exemplo, utiliza um sistema de etiquetas de classificação para as frases as quais analisa (AGÊNCIA LUPA, 2015). As sete etiquetas em questão são: verdadeiro, falso, falta contexto, exagerado, subestimado, contraditório e insustentável. Vale destacar que todo o processo da Agência Lupa não é feito de forma automatizada. A checagem é feita por jornalistas e envolve a análise detalhada de informações e declarações,

buscando fontes confiáveis e verificando a autenticidade dos dados apresentados (AGÊNCIA LUPA, 2015). O processo inclui a seleção de frases relevantes, a busca por evidências em fontes primárias e secundárias, e a classificação da veracidade das informações em diferentes categorias de acordo com as etiquetas (AGÊNCIA LUPA, 2015). Segundo Bentes (2019), essa etiquetagem seria uma tradução da modulação da verdade, deixando de ser algo absoluto, para uma modulação em torno da verdade.

Ademais, Bentes (2019) introduz uma reflexão instigante ao questionar: "Fake News e boatos podem ser desmentidos, mas e as crenças?". Conforme a palestrante, surge o desafio de como se poderia deslegitimar crenças, exceto por meio de uma análise que considere seu impacto ético. Reis et al. (2020) examinaram a disseminação de desinformação em grupos públicos de WhatsApp no Brasil e na Índia, especialmente durante as eleições nacionais recentes nesses países. A pesquisa investiga a extensão com que as imagens classificadas como desinformação por agências de checagem de fatos continuam a ser compartilhadas mesmo após terem sido verificadas como falsas. Ao utilizar dados coletados de grupos públicos de WhatsApp e verificações de agências de *fact-checking*, o estudo buscou compreender como o WhatsApp poderia se beneficiar de conteúdo já checado para reduzir a propagação de notícias falsas, propondo até mesmo uma arquitetura para detecção de desinformação que respeita a privacidade dos usuários. A análise dos pesquisadores mostrou que 40,7% das imagens de desinformação no Brasil e 82,2% na Índia continuaram a ser compartilhadas mesmo depois de já terem sido classificadas como falsas pelas agências de checagem de fatos.

Embora a explosão de iniciativas de verificação e desmascaramento de fatos seja admirável, há uma necessidade urgente de entender quais seriam os formatos mais eficazes para despertar curiosidade e ceticismo no público acerca das informações que consomem e as fontes dessas informações. Simplesmente empurrar mais 'informações factuais' para o ecossistema, sem entender suficientemente os elementos emocionais e ritualísticos da comunicação, pode ser um desperdício de tempo e recursos (WARDLE; DERAKHSHAN, 2023, p. 151).

Ribeiro et al. (2019) conduziram um estudo que analisou anúncios ligados à Rússia que foram veiculados no *Facebook* durante o período das eleições presidenciais dos Estados Unidos em 2016. A pesquisa revelou que esses anúncios tinham um caráter altamente divisivo e eram direcionados principalmente a grupos com base em raça e inclinação política. Esses anúncios mostraram-se dez vezes mais eficazes em engajamento do que um anúncio típico do Facebook, indicando uma capacidade significativa de influenciar os usuários. Além disso, o estudo destacou que o direcionamento desses anúncios tendia a focar em usuários que eram menos propensos a identificar sua impropriedade, evidenciando o uso estratégico das

ferramentas de segmentação da plataforma para atingir audiências mais vulneráveis à desinformação (RIBEIRO et al., 2019).

Figura 2 - Anúncio político no *Facebook* personificando candidatos políticos a figuras religiosas.



97 Reactions · 15 Comments · 29 Shares

Fonte: BENEVENTO, 2024.

Figura 3 - Postagem de política no X, antigo Twitter, comparando eleições à luta do bem contra o mal.



13Mil 20Mil 21Mil

Fonte: LEITÃO, 2021.

As figuras 2 e 3 evidenciam a forte presença da polarização política associada a discursos religiosos nas campanhas eleitorais dos Estados Unidos e do Brasil. Na figura 2, o anúncio patrocinado publicado por uma página de *Facebook* intitulada "Army of Jesus"¹⁵ utiliza referências religiosas para enquadrar a disputa eleitoral entre Hillary Clinton e Donald Trump como uma batalha entre o "bem" e o "mal", personificando Clinton como Satan e Trump como Jesus. O texto da imagem diz: "SATAN: SE EU GANHAR, CLINTON GANHA! JESUS: NÃO SE EU PUDER EVITAR! CLIQUE EM 'CURTIR' PARA AJUDAR JESUS A GANHAR!" Tal narrativa, com um apelo emocional e religioso, sugere que o apoio a Trump é moralmente correto e divinamente aprovado, uma tática que visa mobilizar eleitores por meio de crenças profundas.

De forma semelhante, a figura 3 mostra uma postagem da deputada federal Carla Zambelli no X, antigo Twitter, adotando um discurso semelhante no contexto brasileiro, ao descrever a disputa presidencial de 2022 como uma "luta do bem contra o mal". A publicação rapidamente repercutiu nas redes sociais, demonstrando como a polarização política é reforçada por narrativas maniqueistas que simplificam questões complexas em categorias absolutas de "bem" e "mal". Ambas as figuras exemplificam como o uso de discursos

¹⁵ Exército de Jesus.

religiosos e a personificação dos candidatos em termos polarizados se tornam ferramentas poderosas de desinformação, influenciando percepções e opiniões de maneira emocional e não necessariamente factual.

Ter que aceitar informações que desafiam nosso senso de identidade pode ser chocante. Independentemente de quão persuasiva uma mensagem possa parecer para um observador neutro, é mais fácil ignorar ou resistir a informações que se opõem à nossa própria visão de mundo. É certo que as evidências sugerem que as verificações de fatos tendem a conduzir o conhecimento dos indivíduos na direção da informação correta, mas certamente isso não é suficiente para substituir completamente a informação falsa ou a desinformação (WARDLE; DERAKHSHAN, 2023, p. 59).

Em consonância com essa dinâmica, a pesquisa conduzida por Segurado et al. (2022) analisou a percepção dos eleitores de São Paulo sobre memes desinformativos nas eleições municipais de 2020. Os autores observaram que, entre os participantes com perfil conservador, houve uma tendência significativa de acreditar nas informações falsas apresentadas ou, pelo menos, de ficar em dúvida sobre sua veracidade. Esses conservadores foram negativamente impactados pelas notícias fraudulentas relacionadas a tópicos sensíveis, como a pandemia e as urnas eletrônicas, uma vez corroboraram o argumento de que a desinformação pode influenciar negativamente o debate democrático e os processos eleitorais (Segurado et al., 2022).

A expansão das tecnologias digitais, inicialmente vistas como um meio de democratizar a comunicação, tem revelado contradições significativas entre o potencial democrático e a segregação social (MAIA; URA, 2022). Conforme discutido por Maia e Ura (2022), o problema não se limita ao acesso desigual à *internet*, mas está profundamente enraizado na coleta e comercialização dos dados dos usuários. O controle da comunicação digital por grandes corporações direciona o *design* dos aplicativos para maximizar o valor econômico através da manipulação de comportamentos, caracterizando o que Zuboff (2018) chama de "capitalismo de vigilância" (MAIA; URA, 2022). Esse modelo redefine as interações digitais, promovendo a adaptação dos usuários a uma sociedade desregulamentada, onde os laços de solidariedade são enfraquecidos e a performance individual se torna um critério central de inserção social (MAIA; URA, 2022).

A concentração de poder nas mãos de poucas corporações, aliada à comercialização dos dados, impõe sérias consequências à esfera pública e às práticas democráticas. Maia e Ura (2022) destacam que, ao priorizar conteúdos sensacionalistas e espetaculares que geram mais engajamento, as plataformas digitais favorecem a disseminação de mensagens de ódio e medo, minando o diálogo democrático. Dessa forma, o que deveria ser um espaço para o debate plural torna-se uma "zona de contágio paranoico", como descrito por Adorno (1992),

onde as mensagens são moldadas para maximizar o lucro e não para promover a verdade ou a coesão social.

O modelo de negócios da Big Tech funciona de tal maneira que deixa de ser relevante se as mensagens disseminadas são verdadeiras ou falsas. Tudo o que importa é se elas viralizam (ou seja, se geram números recorde de cliques e curtidas), uma vez que é pela análise de nossos cliques e curtidas, depurados em retratos sintéticos de nossa personalidade, que essas empresas produzem seus enormes lucros. Verdade é o que gera mais visualizações. Sob a ótica das plataformas digitais, as fake news são apenas as notícias mais lucrativas (Morozov, 2018: 11).

De acordo com o OECD Truth Quest Survey (2024), os brasileiros apresentaram o pior desempenho na identificação da veracidade de notícias online entre os 21 países avaliados. Com uma taxa de acerto de apenas 54%, o Brasil ficou significativamente abaixo da média global de 60%. Isso indica que os adultos brasileiros enfrentam maiores dificuldades em distinguir entre conteúdo factual e informações enganosas *online*, em comparação com países como Finlândia, que obteve a maior pontuação com 66%. Esses resultados evidenciam a necessidade de maior investimento em estratégias de alfabetização midiática no Brasil para aprimorar a habilidade da população em reconhecer desinformação.

Em suma, o fenômeno da desinformação representa um desafio complexo que vai além da simples classificação binária do que é verdadeiro ou falso. Como demonstrado, o contexto das informações, as crenças individuais e a maneira como as notícias fraudulentas são estrategicamente direcionadas indicam que a questão demanda não apenas verificação de fatos, mas também uma abordagem mais ampla que considere os aspectos sociais e éticos envolvidos. O combate à desinformação, portanto, requer soluções que integrem múltiplas perspectivas e estratégias, reforçando a necessidade de um entendimento mais profundo sobre o papel da verdade em nossa sociedade contemporânea.

3 CAMINHOS PARA ANÁLISE DA VERIFICAÇÃO AUTOMATIZADA DE NOTÍCIAS FRAUDULENTAS

Para alcançar o objetivo de examinar o uso de abordagens automatizadas na contenção da desinformação que possam ser utilizadas por estudantes, a pesquisa será conduzida de forma qualitativa e exploratória. Minayo (2001, p. 21-22) define a pesquisa qualitativa como a que “trabalha com significados, aspirações, crenças, valores e atitudes, o que corresponde a um espaço mais profundo das relações, dos processos e dos fenômenos que não podem ser reduzidos à operacionalização de variáveis”.

O *corpus* de análise será composto por notícias previamente identificadas como potencialmente fraudulentas, coletadas principalmente de portais de notícias e redes sociais.

Inicialmente, as notícias serão verificadas por meio de duas ferramentas automatizadas de checagem: o *bot* Fátima 3.0, da agência de *fact-checking* Aos Fatos e a plataforma de detecção de notícias fraudulentas FakeCheck, desenvolvida no projeto "Detecção Automática de Notícias Falsas para o Português" financiado pelo PIBIC¹⁶ do CNPq e também pela CAPES. Ambos, o *bot* Fátima 3.0 e a plataforma FakeCheck, foram desenvolvidos para operar especificamente em língua portuguesa.

Caso a ferramenta Fátima 3.0 não forneça respostas satisfatórias ou completas, será feita uma segunda etapa de análise utilizando métodos tradicionais de checagem, que incluem a comparação com fontes confiáveis, a análise do contexto e a avaliação de informações por agências de *fact-checking* humanas. Essa abordagem híbrida permitirá avaliar a eficácia e praticidade das soluções automatizadas na contenção da desinformação e como essas ferramentas podem ser aplicadas no contexto educacional, contribuindo para a formação crítica dos estudantes. Neste caso somente considerei esta ferramenta em específico, pois ela não fornece respostas dicotômicas.

Como forma de delimitar o *corpus* da pesquisa, foram selecionadas notícias e afirmações de cunho político com grande visibilidade no cenário brasileiro, visto que, devido ao período do ano, as propagandas e a desinformação políticas estão em alta, influenciando diretamente a opinião pública e o processo democrático. Além disso, considero apenas notícias e afirmações dentro do período específico de abril de 2024 a outubro de 2024, ou seja, cerca dos últimos seis meses. No que diz respeito à segunda etapa do processo de verificação, ela será baseada no guia desenvolvido pela FactBar EDU: "*Fact-checking for Educators and Future Voters*" (2018). Em específico, o tópico 2, ajustando para a realidade da pesquisa em questão. Em suma, esse guia da FactBar EDU adaptou o método de *fact-checking* para o ambiente educacional, propondo etapas que incluem a seleção de uma afirmação, o exame da afirmação utilizando diferentes fontes, a elaboração de um relatório de verificação com as descobertas e a apresentação dos resultados para discussão.

¹⁶ Programa Institucional de Bolsas de Iniciação Científica.

4 À LUZ DO *CORPUS*: ANÁLISE E DISCUSSÃO

Antes de tudo, se faz necessário destacar que o *corpus* utilizado nesta pesquisa, de notícias previamente identificadas como fraudulentas e de grande incidência nas redes, no cenário brasileiro, apesar de adequado para o escopo exploratório do estudo, não é extenso o suficiente para garantir uma alta precisão ou generalização dos resultados. Devido ao número limitado de notícias analisadas, os achados devem ser interpretados com cautela, pois não refletem necessariamente uma medida definitiva da eficácia dessas ferramentas. A precisão e resultados dos bots pode variar em um corpus maior e mais diversificado, portanto, os achados apresentados aqui fornecem apenas uma visão inicial e exploratória, que pode ser aprofundada com estudos futuros utilizando um volume maior de dados.

Tabela 2 - Relação do *corpus* de notícias analisadas

Chamada da notícia	Fonte	Resultado Fátima 3.0	Observações	Resultados FakeCheck	
				1*	2*
Pablo Marçal é diagnosticado com traumatismo no tórax	Portal de notícia	Parcialmente verdadeira	Respondeu diretamente ao enviar o <i>link</i> da notícia. Resposta satisfatória.	Parece verdadeira.	Parece verdadeira.
Golpe de 1964 fez do STF um enfeite institucional no Brasil	Portal de notícia	Enganosa	Informou que não tinha informações suficientes para responder à pergunta. Ao contextualizar, não forneceu uma resposta satisfatória. Ao questionar diretamente "STF já foi cassado?", o <i>bot</i> forneceu uma resposta satisfatória.	Parece verdadeira.	Parece verdadeira.
Pesquisa Quaest aponta derrocada de Boulos e fora do 2º turno	Portal de notícia	Enganosa	Informou que não tinha informações suficientes para responder à pergunta. Mas ao contextualizar forneceu uma resposta satisfatória.	Parece verdadeira.	Pode ser falsa.
Lula diz que conseguiu acabar com a fome em 2014	Portal de notícia	Exagerada	Respondeu diretamente ao enviar o <i>link</i> da notícia. Resposta satisfatória.	Parece verdadeira.	Parece verdadeira.
Relatório das Nações Unidas não identifica fome no norte de Gaza	Portal de notícia	Inconclusiva	Não soube responder diretamente e não deu contexto. Verificação manual foi necessária.	Parece verdadeira.	Parece verdadeira.
Coreia do Norte	Portal de	Inconclusiva	Não soube responder	Parece	Parece

proíbe karaokê, dizendo que “cheira” a cultura “capitalista podre”	notícia		diretamente e não deu contexto. Verificação manual foi necessária.	verdadeira.	verdadeira.
Como Kamala Harris e Joe Biden estão financiando a censura no Brasil	Portal de notícia	Enganosa	Informou que não tinha informações suficientes para responder à pergunta. Mas ao contextualizar forneceu uma resposta satisfatória.	Parece verdadeira.	Parece verdadeira.
Mais uma eleição sob censura	Portal de notícia	Enganosa	Respondeu diretamente ao enviar o <i>link</i> da notícia. Resposta satisfatória.	Pode ser falsa.	Parece verdadeira.
Governo Lula suspende 125 mil auxílios-doença e prepara mutirão para atualização de dados	Portal de notícia	Inconclusiva	Informou que não tinha informações suficientes para responder à pergunta. Ao contextualizar, não forneceu uma resposta satisfatória. Verificação manual foi necessária.	Parece verdadeira.	Parece verdadeira.
Projeto para confiscar bilhões dos brasileiros na “calada da noite” é obstruído na câmara	Instagram	Enganosa	Não soube responder diretamente e não deu contexto. Ao enviar a chamada da notícia, informou que não sabia responder diretamente à pergunta. Mas ao contextualizar retornou uma resposta satisfatória.	Pode ser falsa.	Pode ser falsa.
ALEXANDRE DE MORAES E EX-FUNCIONÁRIOS DO X INTERFERIRAM NAS ELEIÇÕES BRASILEIRAS DE 2022, DIZ ELON MUSK	Instagram	Enganosa	Informou que não tinha informações suficientes para responder à pergunta. Mas ao contextualizar forneceu uma resposta satisfatória.	Pode ser falsa.	Pode ser falsa.
BRASIL SE UNE A COREIA DO NORTE E PROÍBE O X/TWITTER POR ORDEM DE ALEXANDRE DE MORAES	Instagram	Inconclusivo	Não soube responder diretamente e não deu contexto. Verificação manual foi necessária.	Pode ser falsa.	Pode ser falsa.
Moraes interferiu nas eleições no Brasil	Facebook	Enganoso	Não soube responder diretamente e não deu contexto. Ao enviar o trecho inicial da notícia retornou uma resposta	Parece verdadeira.	Parece verdadeira.

			satisfatória.		
Nos últimos dias, todos os brasileiros testemunharam uma série de acontecimentos... [Texto 1]	Telegram	Inconclusiva	Não soube responder diretamente e não deu contexto. Verificação manual foi necessária.	Parece verdadeira.	Parece verdadeira.
“Devolveremos este território e construiremos a Grande Khazaria [...]”... [Texto 2]	Telegram	Inconclusiva	Não soube responder diretamente e não deu contexto. Verificação manual foi necessária.	Pode ser falsa.	Pode ser falsa.
Isso também comprova que o Brasil desde 1500 nunca foi um país Livre e muito menos Democrático... [Texto 3]	Telegram	Inconclusiva	Não soube responder diretamente e não deu contexto. Verificação manual foi necessária.	Pode ser falsa.	Parece verdadeira.
Parece que os Satanistas estão conseguindo mesmo retirar e destruir tudo o que faz o povo se lembrar do Governo dos Militares... [Texto 4]	Telegram	Inconclusiva	Não soube responder diretamente e não deu contexto. Verificação manual foi necessária.	Pode ser falsa.	Parece verdadeira.

Fonte: A autora (2024)

Tabela 3 - Textos retirados do Telegram

Texto 1 do Telegram	Texto 2 do Telegram
<p>– Nos últimos dias, todos os brasileiros testemunharam uma série de acontecimentos que lançaram ainda mais luz sobre os gravíssimos retrocessos que a liberdade e a segurança jurídica têm sofrido no Brasil.</p> <p>– O X foi banido por questionar decisões judiciais que exigiam não apenas a remoção pontual de postagens, mas a exclusão permanente de perfis e páginas. Isso configura censura prévia, uma prática claramente vedada pela nossa Constituição e uma grave violação de direitos fundamentais.</p> <p>– Pior: ao banir a principal rede social do Brasil, punimos não uma empresa ou um empresário, mas milhões de brasileiros, que ficam privados de acesso à informação livre e impedidos de expressar o que pensam.</p> <p>– Pouco antes dessa decisão, o mesmo ministro que baniu o X censurou a Folha de São Paulo, impedindo um dos maiores jornais do país de publicar uma entrevista com Filipe Martins, um inocente que ficou preso por</p>	<p>“Devolveremos este território e construiremos a Grande Khazaria – o estado judeu – nestas terras férteis da mesma forma que, há 50 anos, criamos Israel.”</p> <p>- Mashiah, Menachem Mendel Schneerson, 1994.</p> <p>O Rebe e Mashiah vivo da Máfia Chabad foi o líder do povo judeu de 1951 a 1994. Ele é considerado a figura judaica mais influente por séculos. Schneerson considerou mais do que Jerusalém e o Grande Israel como uma pátria ancestral; o Rebe tinha uma forte afinidade com suas raízes Ashkenazi, pois nasceu na Ucrânia em 1902. Ele desejava retornar a esta 'Terra Prometida' para atuar como a capital global do Reino dos Judeus.</p> <p>Em um discurso de 1994 publicado na publicação Slavyanin, de Vologda, Menachem Mendel Schneerson detalhou suas intenções de destruir tanto a Ucrânia quanto a Rússia.</p> <p>“ Os eslavos, entre eles os russos – são o povo mais inflexível do mundo. Os eslavos são inflexíveis como</p>

mais de seis meses sem julgamento sob a alegação de que ele havia feito uma viagem que ele nunca fez.

– O mesmo ministro transformou uma banal discussão por causa de uma fila de aeroporto em um "atentado contra a democracia", tornando uma família ordeira e cumpridora da lei em alvo de um inquérito policial sem qualquer justificativa legal.

– Agora, descobrimos que a Starlink também está sendo punida arbitrariamente, tornando-se vítima colateral de uma disputa judicial com a qual nada tem a ver. Isso mina ainda mais a segurança jurídica no Brasil e a confiança dos investidores internacionais no país. O Brasil precisa ser um porto seguro para o investimento e a inovação, mas infelizmente temos caminhado na direção oposta, afugentando investidores e rejeitando empresas que estão entre as mais inovadoras do mundo. Minar essa confiança coloca em risco nossa capacidade de crescer, de gerar empregos e de trazer novas tecnologias que impulsionam o desenvolvimento do país.

– Tudo isso vem à tona em um momento em que, dia após dias, novas revelações do jornalista Glenn Greenwald, publicadas pela Folha de São Paulo, nos mostram que esses inquéritos são conduzidos sem qualquer respeito pelo devido processo legal e com a pré-seleção de alvos. Nesses inquéritos, as pessoas não estão sendo investigadas com base em fatos ou por suas ações, mas em decorrência de quem elas são e do que pensam e defendem politicamente.

– Quando a justiça age seletivamente desse modo, todos corremos o risco de perder as garantias que nos são asseguradas pela Constituição. Quando a censura prévia é normalizada, toda a sociedade perde sua liberdade sua dignidade. Quando a liberdade de expressão e a liberdade de imprensa são ameaçadas, a Democracia passa a gritar por socorro.

– Por tudo isso, faço um chamado a todos os brasileiros que amam a liberdade e se preocupam com a saúde da nossa democracia: compareçam na Avenida Paulista no próximo sábado, dia 7 de setembro!

– A defesa de tudo aquilo que torna uma nação livre, justa e democrática é responsabilidade de todos nós, independentemente de posicionamento político ou preferências eleitorais.

– Se você está em São Paulo ou pode ir para lá neste final de semana, não deixe para acompanhar as manifestações apenas pela TV ou pela Internet. Junte-se a nós, pessoalmente, na Avenida Paulista!

– Sábado, 07 de setembro de 2024, a liberdade precisa de todos nós.

resultado de suas habilidades psicológicas e intelectuais, criadas por muitas gerações de ancestrais. É impossível alterar esses genes. Eslavo, russo, pode ser destruído, mas nunca conquistado.

É por isso que essa semente está sujeita à liquidação e, a princípio, a uma redução acentuada em seus números. Primeiro, dividiremos as nações eslavas em pequenos países com conexões fracas e rompidas. Para isso, usaremos nosso antigo método: Dividir e conquistar. Tentaremos colocar esses países uns contra os outros e sugá-los para guerras civis em prol da destruição mútua. Os ucranianos pensariam que estão lutando contra a Rússia expansionista e lutando por sua independência. Eles pensarão que finalmente ganharam sua liberdade, enquanto se tornam totalmente subjugados por nós. O mesmo será pensado pelos russos, como se defendessem seus interesses nacionais para devolver suas terras, "ilegalmente" tiradas deles, e assim por diante." Olhando para trás na história, deve-se admitir que essas terras são as antigas terras ancestrais da Khazaria judaica, isto é, Israel, capturada pela Rus' de Kiev (o antigo estado da Rússia com capital em Kiev) no século X.

Os eslavos são hóspedes temporários nessas terras e estão sujeitos a despejo. Nós devolveremos esse território e construiremos a Grande Khazaria – o estado judeu – nessas terras férteis da mesma forma que, 50 anos atrás, criamos Israel, expulsando os palestinos. Os israelenses se mudarão parcialmente para cá, e nós expulsaremos o gado eslavo para bem ao norte, além de Moscou. Haverá um pequeno Território do Norte, uma reserva com uma população compacta — uma reserva, como as reservas indígenas na América..."

Schneerson discute um império Khazar ressurgente ; o Chabad Lubavitch pretende se mudar de Nova York para Dnepropetrovsk , desconsiderando as escrituras e profecias do sionismo e de Israel. Os Lubavitchers mantêm uma visão escritural bastarda, reconhecendo Menachem Mendel Schneerson, que faleceu em Nova York em 1994, como Mashiach.

Fonte: Conta do Telegram do ex-presidente da República Jair Messias Bolsonaro (@jairbolsonarobrasil)

Texto 3 do Telegram

30/09/2023 - Isso também comprova que o Brasil desde 1500 nunca foi um país Livre e muito menos Democrático como os do G7 e especialmente como os EUA. Nestas cédulas tem mais de 100 tipos de escolhas para os Americanos fazerem como até escolher via

Fonte: Grupo do Telegram

Texto 4 do Telegram

03/10/2024 - Parece que os Satanistas estão conseguindo mesmo retirar e destruir tudo o que faz o povo se lembrar do Governo dos Militares que tinha problemas mas nunca foi uma Ditadura de verdade onde muitos idiotas ainda insistem em chamar de "Ditadura" com a

eleições sem fraudes é claro se querem que as drogas sejam liberadas ou não no nível federal, estadual e municipal. Isso é uma das coisas que os Jornazistas e a Imprensalha esconderam da população Brasileira por milhares de anos a fio.

Iago Barbosa, brasileiro correspondente nos EUA, mostra para a audiência da Auri Verde a cédula para votação presidencial dos Estados Unidos. Iago recebeu esse documento mesmo não sendo oficialmente uma cidadão americano no momento.

desculpa de que aprenderam isso na "escola" e que aceitam a mentira da verdade oficial que foi colocada nos livros didáticos por Terroristas brutais, sanguinários, subversivos e ultra perigosos que são os Socialista-Comuno-Nazi-Fascista-Feudalista-Totalitario-Autoritario-Genocida-Escravagista-Infanticida-Satanista a partir de 1985. A única coisa que os militares não deixavam era a imoralidade satanica tomar conta da mídia e da imprensa isso era regulado mas o debate político era livre sem censura e não havia perseguição para quem como o cantor Raul Seixas que se declarava um anarquista ou mesmo o Cazuza que dizia ser da esquerda social como os Kennedys fossem censurados e punidos pelo governo dos militares ou seja os Militares aplicavam de forma correta o princípio de Liberdade de Expressão que a Constituição dos EUA da maneira exata e objetiva como foi escrita a famosa Primeira Emenda da Constituição dos EUA no Brasil de 1964 até 1985. A morte do Cid Moreira me deixou muito mais triste do que a do próprio Silvio Santos. Sei que o já saudoso Cid Moreira era de um outro tempo onde o povo do Ocidente não só do Brasil eram muito mais temente à Deus e ao Senhor Jesus Cristo. Cid Moreira era de um tempo onde o povo se submetida à moral Cristã e que obedecia mais à Lei e o padrão moral de Deus. Que saudades do tempo em que a Rede Globo era a TV da Família Brasileira assim como o SBT, Record e das extintas TV Manchete, Excelsior, Gazeta e a Tupi do Assis Chateaubriand. Gideões: Morre Cid Moreira, aos 97 anos, de pneumonia. Cid foi o primeiro apresentador do Jornal Nacional e também ficou muito conhecido por fazer a narração da Bíblia. Após os comunistas assumirem o poder em 86, Cid perdeu suas posições na TV porque tinha um perfil cristão e conservador. Esta é a verdade.

Fonte: Grupo do Telegram

Fonte: Grupo do Telegram

Fonte: A autora (2024)

Para classificar os resultados fornecidos pelo *bot* Fátima 3.0, adotei uma abordagem que considera a diversidade de respostas geradas pela ferramenta. Como o *bot* não se limita a classificações binárias como "verdadeiro" ou "falso", as análises foram organizadas em categorias que refletem melhor a complexidade dos resultados: "parcialmente verdadeira", "descontextualizada", "enganosa" ou "inconclusiva". Essa metodologia permitiu capturar nuances nas verificações e proporcionar uma interpretação mais adequada ao contexto de cada notícia. Além disso, deve-se enfatizar que o propósito não é comparar as ferramentas entre si, mas a eficácia de cada uma delas. No que se refere ao FakeCheck, a subdivisão na coluna de resultados da ferramenta refere-se ao modelo de detecção, onde 1 diz respeito ao modo de classificação Palavras do Texto e 2 ao de Classes Gramaticais.

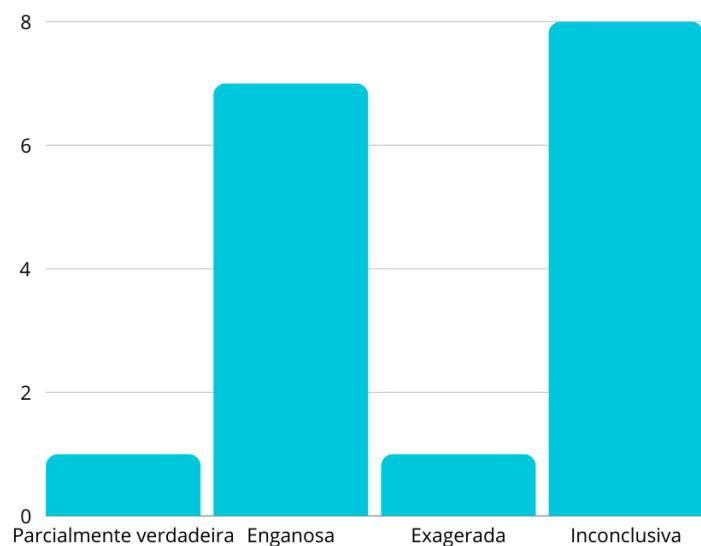
Como mencionado na tabela 1, o FakeCheck aplica modelos de aprendizado de máquina para classificar notícias como "verdadeiras" ou "falsas". Além disso, os modelos de classificação do detector foram treinados com o *corpus* Fake.Br, o qual o funcionamento foi discutido na seção 2.1. Em síntese, o sistema funciona extraíndo atributos linguísticos do texto, como palavras e classes gramaticais, usando representações como "*Bag of Words*"¹⁷ e análises gramaticais para detectar padrões típicos de desinformação. Esses atributos são aplicados em um classificador *Support Vector Machine* (SVM)¹⁸ que infere a veracidade do conteúdo.

Ao todo o *corpus* foi composto por 17 notícias retiradas de diferentes veículos de comunicação, como ilustrado na tabela 2, sendo a maioria, 9 no total, retirada de portais de notícias. As outras 8 notícias foram retiradas das redes sociais: 4 do Telegram (ver tabela 3), 3 do Instagram e 1 do Facebook.

¹⁷ É uma técnica de representação de texto em que uma frase, documento ou *corpus* é transformado em uma coleção de palavras sem considerar a ordem ou estrutura gramatical.

¹⁸ Trata-se de um algoritmo de aprendizado supervisionado usado principalmente para classificação e regressão.

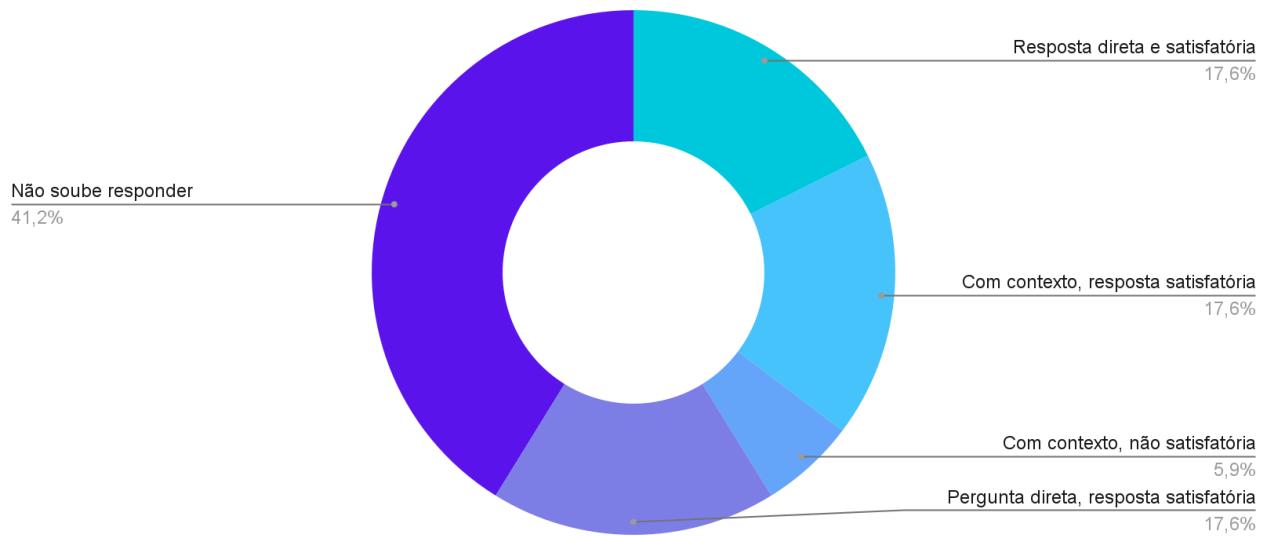
Gráfico 1 - Resultado de classificação de notícias do *bot* Fátima 3.0



Fonte: A autora (2024)

A análise dos resultados do *bot* Fátima 3.0, conforme mostrado na tabela 2, revela uma classificação diversa das notícias verificadas. Entre as 17 notícias analisadas, o *bot* apresentou as seguintes classificações: parcialmente verdadeira, 1 notícia; enganosa, 7 notícias; exagerada, 1 notícia; inconclusiva, 8 notícias (ver gráfico 1). Das 8 notícias classificadas como enganosas, 5 eram oriundas de portais de notícias e as outras 3 de redes sociais (2 do Instagram e 1 do Facebook). É possível dividir o desempenho das respostas de Fátima 3.0 em 5 grupos: Teve resposta direta e consequentemente, resposta satisfatória; Não soube responder, mas com contexto obteve resposta satisfatória; Não soube responder e com contexto não obteve resposta satisfatória; Não soube responder, mas ao perguntar diretamente teve resposta satisfatória; Não soube responder, não teve resposta satisfatória e precisou de verificação manual (ver gráfico 2).

Gráfico 2 - Grupos de classificação do desempenho de respostas de Fátima 3.0



Fonte: A autora (2024)

Cerca de metade das classificações fornecidas foi inconclusiva. Nesses casos, o *bot* não soube responder diretamente à consulta ou não conseguiu contextualizar adequadamente as informações, nem mesmo com uma pergunta direta, exigindo verificação manual. Isso ocorreu com frequência em notícias compartilhadas em redes sociais (Telegram e Instagram), o que sugere uma dificuldade maior do *bot* em processar conteúdos menos estruturados ou com menor transparência de fonte. Em contraste, para as notícias originadas de portais de notícias, como a informação de "Pablo Marçal diagnosticado com traumatismo no tórax", o *bot* foi capaz de responder diretamente, proporcionando respostas satisfatórias.

Notas relevantes mostram que as notícias originadas de redes sociais, especialmente Telegram, foram majoritariamente classificadas como inconclusivas, enquanto notícias vindas de portais de notícias tradicionais obtiveram classificações mais precisas (enganosa ou parcialmente verdadeira). Isso pode indicar que a ferramenta tem melhor desempenho ao lidar com conteúdos mais estruturados e verificáveis. Convém salientar que as notícias selecionadas tiveram grande repercussão e consequentemente tiveram mais atenção das agências de verificação. O caso provavelmente seria diferente caso se tratasse de notícias com repercussão mais local, em regiões específicas, e sem grande comoção nacional.

Ao analisar os resultados obtidos pela ferramenta FakeCheck na tabela 2, observou-se que a detecção de notícias fraudulentas por meio dos dois modelos disponíveis apresentou

certas variações. As 17 notícias verificadas foram analisadas com base nos modelos de "Palavras do Texto" e "Classes Gramaticais". No primeiro modelo, que considera palavras do texto, 7 das 17 notícias foram classificadas como "Pode ser falsa". Já no segundo modelo, que utiliza classes gramaticais, 5 das 17 notícias receberam essa mesma classificação. Somente em 4 casos, dos 8 casos ao todo classificados como falsos, houve concordância entre os dois modelos de classificação.

As notícias classificadas como inconclusivas pelo Fátima 3.0 — provenientes principalmente de redes sociais como Telegram e Instagram — foram as que o FakeCheck mais identificou como potencialmente falsas. Essas notícias frequentemente apresentam erros ortográficos e utilizam termos e expressões carregadas, como "Socialista-Comuno-Nazi-Fascista-Feudalista-Totalitário-Autoritário-Genocida-Escravagista-Infanticida-Satanista" ou "onde muitos idiotas ainda insistem em chamar de 'Ditadura' com a desculpa de que aprenderam isso na 'escola'". Além disso, o uso de uma linguagem sensacionalista, com expressões como "na calada da noite" e "expropriação de R\$ 8 bilhões esquecidos em contas bancárias", reforça essa tendência. Como observado no estudo de Monteiro et al. (2018), características como a pausalidade, emotividade, incerteza e não-imediatismo são fundamentais na análise linguística dessas mensagens. A partir da análise dessas características, o FakeCheck consegue detectar padrões presentes em textos fraudulentos, classificando-os como falsos. Por exemplo, o uso excessivo de adjetivos e advérbios (emotividade) ou a presença de pronomes pessoais (não-imediatismo) em uma notícia pode indicar um viés ou uma tentativa de manipulação emocional, o que contribui para a classificação da notícia como "falsa".

Uma explicação para o melhor desempenho do FakeCheck com notícias provenientes de redes sociais, comparado a portais de notícias, pode estar no fato de que qualquer pessoa pode mais facilmente ser autora de conteúdo nessas plataformas, as quais têm certas características de linguagem e o uso de certos padrões linguísticos e gramaticais como toda mídia. Nas redes sociais, os textos costumam ser mais informais e muitas vezes carregados de emoções. Essas características, como o uso excessivo de adjetivos, simplificação de informações e falta de uma estrutura factual clara, são identificadas com maior facilidade pelos modelos de aprendizado de máquina. Por outro lado, os portais de notícias tradicionais tendem a seguir padrões jornalísticos mais rígidos e estruturados, o que pode confundir o FakeCheck ao tentar identificar desinformação, uma vez que as notícias fraudulentas nesses portais são geralmente mais sofisticadas e difíceis de detectar com base em características linguísticas.

Tabela 4 - Relação das notícias verificadas manualmente

Chamada da notícia	Fonte	Comentários	Resultado
Relatório das Nações Unidas não identifica fome no norte de Gaza	Portal de notícia	No corpo da própria notícia há um link para o relatório divulgado pelas Nações Unidas. O Relatório do Comitê de Revisão da Fome (FRC) não confirma a classificação de fome para o norte de Gaza devido à falta de dados confiáveis. A afirmação de que não identifica fome no norte de Gaza é equivocada.	Enganosa
Coreia do Norte proíbe karaokê, dizendo que “cheira” a cultura “capitalista podre”	Portal de notícia	Ao buscar no Google pela afirmação de que a Coreia do Norte proibiu o karaokê, não houve nenhum resultado. Ao procurar pela mesma frase em inglês, uma notícia correspondente foi achada. A notícia em português é nada mais nada menos que uma tradução da notícia em inglês. Nenhuma fonte é mencionada, o que sugere plágio. Nenhuma fonte confiável menciona que o karaokê foi proibido no país.	Enganosa
Governo Lula suspende 125 mil auxílios-doença e prepara mutirão para atualização de dados	Portal de notícia	Ao pesquisar pela notícia no Google, um dos primeiros resultados foi o da agência de checagem Lupa, a qual classificou como "falta contexto".	Falta contexto
BRASIL SE UNE A COREIA DO NORTE E PROÍBE O X/TWITTER POR ORDEM DE ALEXANDRE DE MORAES	Instagram	Para falar sobre a suspensão do X no Brasil, o autor liga diretamente o país à Coreia do Norte, onde a rede social não circula. A priori, não há relação alguma entre as razões pelas quais o X foi suspenso em ambos os países. A própria publicação explica que o site foi bloqueado por não nomear um representante legal no país. No site do STF é possível entender melhor a situação. É possível afirmar que a notícia em questão publicada no Instagram tem a intenção de causar pânico moral.	Enganosa
Nos últimos dias, todos os brasileiros testemunharam uma série de acontecimentos...	Telegram	Alegar que a remoção de perfis em plataformas constitui "censura prévia" sem contextualizar o motivo da remoção pode ser enganoso, já que a Constituição veda a censura prévia pelo Estado, mas plataformas digitais seguem regulamentações diferentes. Além disso, o texto menciona que milhões de brasileiros foram afetados. No entanto, banir uma plataforma, mesmo que temporariamente, não elimina o acesso geral à informação, já que há múltiplas fontes. A alegação de que as pessoas são investigadas "por quem são e pelo que pensam" é generalizada e sem apresentação de evidências claras. Isso pode induzir ao erro ao sugerir que as	Enganosa e falta contexto.

		investigações são feitas de forma seletiva e injusta, sem detalhar os fatos jurídicos por trás de cada caso mencionado.	
“Devolveremos este território e construiremos a Grande Khazaria [...]”...	Telegram	O texto nega que o período militar no Brasil (1964-1985) tenha sido uma ditadura, o que contradiz historiadores e documentos oficiais que confirmam a repressão política, censura e violência estatal. A referência a uma suposta conspiração de "terroristas" e "Satanistas" é infundada e carece de evidências. Esses termos são usados de maneira exagerada e sem fundamento factual. O texto também afirma que não havia censura durante o regime militar, o que é historicamente incorreto, já que havia forte controle da mídia e repressão a opositores.	Enganosa
Isso também comprova que o Brasil desde 1500 nunca foi um país Livre e muito menos Democrático...	Telegram	A alegação de que o Brasil nunca foi um país livre ou democrático desde 1500 é exagerada. O Brasil passou por diversos períodos políticos, incluindo a ditadura e a redemocratização, mas afirmar que nunca teve democracia é incorreto, especialmente considerando que desde 1985 vive em um regime democrático. Também, a ideia de que a imprensa "escondeu" algo dos brasileiros por "milhares de anos" é imprecisa, já que não há evidências de que a mídia esconde informações desse nível.	Enganosa
Parece que os Satanistas estão conseguindo mesmo retirar e destruir tudo o que faz o povo se lembrar do Governo dos Militares...	Telegram	Aqui cabe a mesma análise do texto 2 do Telegram acerca da ditadura militar. Expressões exageradas e sem base, como "terroristas brutais", "Satanistas", e a mistura de termos como "Socialista-Comunista-Nazi-Fascista", servem para confundir e criar uma narrativa sem fundamentos históricos.	Enganosa e sem fundamento

Fonte: A autora (2024)

No que diz respeito às notícias que não obtiveram um retorno satisfatório das ferramentas analisadas em questão, foi preciso buscar e analisar mais a fundo. Por conseguinte, um ponto importante a ser mobilizado na análise do *corpus* é a articulação multidisciplinar no combate à desinformação. Pires (1998) argumenta que a fragmentação do ensino dificulta a formação crítica dos alunos e sugere uma integração entre diferentes áreas do conhecimento. Assim, tomando como base o guia da FactBar EDU, as atividades de checagem de fatos foram estruturadas para tornar o processo de verificação acessível ao ambiente escolar, facilitando sua aplicação por alunos. O método sugerido envolve cinco etapas principais: seleção de uma afirmativa, análise de sua veracidade, verificação dos fatos,

produção de um relatório e compartilhamento dos resultados. Esse processo é adaptado de práticas jornalísticas, mas ajustado para o nível dos estudantes, incentivando-os a investigar de forma crítica e a desenvolver habilidades de pesquisa e argumentação.

A seleção das afirmativas é guiada por critérios que ajudam os alunos a identificar informações ambíguas ou controversas em diversas fontes, como entrevistas, redes sociais ou vídeos *online*. Na sequência, o exame da afirmativa e a verificação das evidências ensinam os estudantes a buscar fontes confiáveis e a questionar a veracidade das informações de forma objetiva. O processo culmina na criação de relatórios que podem ser compartilhados em diversos formatos, reforçando o aprendizado e promovendo a responsabilidade cívica. Essa abordagem não só aprimora as competências de pensamento crítico dos alunos, como também os prepara para lidar com a desinformação no ambiente digital.

Levando em consideração que as notícias já estão estabelecidas e que este é um contexto de análise de ferramentas automatizadas, abordei somente as etapas de análise da veracidade e verificação de fatos. Além disso, ao invés de uma afirmativa em específico, estou analisando uma notícia. Na etapa de análise da veracidade, os alunos são orientados a desmembrar a afirmativa selecionada para torná-la mais clara e comprehensível. Cada parte da declaração é analisada, identificando elementos ambíguos ou passíveis de interpretação. Além disso, essa fase envolve a busca por fontes relevantes e confiáveis que permitam corroborar ou refutar a afirmativa. Os alunos são incentivados a utilizar diversas fontes de informação, como artigos, publicações científicas ou notícias, sempre questionando a confiabilidade dessas fontes. Nesse sentido, a teoria dos multiletramentos, discutida por Rojo (2015), mostra-se fundamental para compreender a necessidade de integrar diferentes formas de linguagem no currículo educacional.

Com base na tabela 4, todas as 8 notícias classificadas como "inconclusivas" apresentaram características que as tornaram difíceis de verificar de maneira definitiva. Desses notícias, 5 eram provenientes de redes sociais (Telegram e Whatsapp), e as demais vieram de portais de notícias. A análise revelou que essas informações eram frequentemente enganosas, carecendo de contexto adequado ou apresentando dados sem fundamentação robusta.

As notícias verificadas foram identificadas como enganosas principalmente por distorcerem fatos ou apresentarem declarações que não correspondiam à realidade quando verificadas em fontes externas. Por exemplo, a afirmação de que "Relatório das Nações Unidas não identifica fome no norte de Gaza" rotulei como enganosa, uma vez que o relatório em questão indicou a falta de dados confiáveis para determinar a situação real, em vez de

negar a presença de fome na região. Da mesma forma, a alegação sobre a proibição de karaokê na Coreia do Norte foi verificada também como enganosa, pois não havia registros confiáveis que sustentam tal proibição, além da notícia se tratar de uma tradução de outro portal e sem fundamentação de fontes oficiais.

Em suma, os resultados sugerem que o *bot* Fátima 3.0 tem dificuldades em processar conteúdos mais informais ou menos diretos, especialmente de redes sociais, que frequentemente compilam mais de um tipo de notícia numa única mensagem ou publicação. Enquanto isso, para portais de notícias tradicionais, ele se mostrou mais eficiente, fornecendo respostas mais conclusivas e contextualizadas. Essa limitação nas redes sociais reflete a necessidade de aprimoramento no processamento de informações de fontes mais ambíguas e com informações mais dispersas. Em contrapartida, o FakeCheck apresentou uma tendência maior a classificar corretamente como falsa, notícias provenientes de fontes menos tradicionais, como as redes sociais. Entretanto, ao analisar notícias de portais a taxa de acerto de classificação como "falsa" foi praticamente zero.

O uso de ferramentas automatizadas no processo de verificação de notícias se mostrou vantajoso ao mirar o uso dessas ferramentas em contextos específicos, como verificação rápida de notícias em *sites* e análise de desinformação em redes sociais. Isso porque agiliza e facilita a análise de grandes volumes de informações as quais somos bombardeados todos os dias. Embora nenhuma tecnologia seja capaz de resolver todos os problemas, essas ferramentas podem tornar os processos mais otimizados. O *bot* Fátima, por exemplo, demonstrou grande eficácia ao verificar informações diretas e *links* de notícias, poupando tempo na busca extensiva por fontes, já que, na maioria dos casos, ele fornece *links* para informações previamente verificadas. Por outro lado, o FakeCheck facilita a verificação do teor de mensagens compartilhadas em redes sociais como Telegram — e também pode-se estender ao WhatsApp, devido à semelhança no tipo de mídia circulante — e Instagram. Contudo, é importante ressaltar que, embora ambas as ferramentas sejam úteis, elas devem ser utilizadas com consciência, uma vez que, em casos mais complexos — que se configuram como a maior parte dos casos — podem exigir uma revisão manual para garantir a precisão dos resultados.

4.1 Mobilização da educação frente à desinformação

O processo de ensino-aprendizagem envolve uma série de fatores interconectados, que vão desde a qualificação do corpo docente até a garantia de infraestrutura adequada, incluindo materiais pedagógicos, equipamentos e ambientes físicos propícios para o aprendizado. Para que a educação seja efetiva, é essencial que essas condições sejam

cuidadosamente estruturadas, permitindo um ambiente de ensino que favoreça tanto os professores quanto os alunos (NETO J. J. S. et al., 2014). Os dados do Censo Escolar de 2011 mostram que 44,5% das escolas brasileiras têm apenas infraestrutura elementar, especialmente nas áreas rurais das regiões Norte e Nordeste, onde atendem mais de 7 milhões de alunos (NETO J. J. S. et al., 2014). Esses números evidenciam a necessidade urgente de investimentos em infraestrutura escolar, conforme analisado por Neto J. J. S. et al. (2013). Estudos descritivos, como os realizados por esses autores, são fundamentais para tornar essa realidade mais clara para os gestores e apoiar a criação de políticas públicas que possam melhorar as condições físicas e de ensino nas escolas brasileiras.

Diante das disparidades socioeconômicas, que há muito tempo caracterizam o cenário do nosso sistema educacional, é fundamental questionar o verdadeiro sentido do "avanço" tecnológico quando aplicado à educação. A incorporação de ferramentas digitais na educação básica vai muito além da simples disponibilização de tecnologias; é necessário refletir sobre a realidade estrutural das escolas. Afinal, quais instituições possuem a infraestrutura adequada para integrar essas inovações de forma eficaz? Muitas escolas, especialmente em regiões menos favorecidas, carecem de recursos básicos como acesso à *internet*, laboratórios de informática ou até mesmo condições físicas adequadas. O debate sobre a inclusão de tecnologias no ensino deve, portanto, considerar as desigualdades estruturais existentes e a capacidade de cada escola de proporcionar um ambiente de aprendizado adequado. Certamente o Brasil, que num retrocesso, está novamente presente no mapa da fome (MACHADO, A. L., 2021), enfrenta adversidades muito diferentes no combate à desinformação em comparação à Finlândia, um país desenvolvido.

Feitas essas considerações, é profícuo pautar mais objetivamente como essas tecnologias podem ser acionadas nas escolas. Para que a educação básica implemente o combate à desinformação, é essencial integrar essa questão no currículo e nas práticas pedagógicas diárias. Projetos de letramento midiático podem ser inseridos no ensino por meio de atividades que envolvam a análise de notícias, uso de plataformas de *fact-checking* e criação de campanhas de conscientização sobre desinformação, não se limitando às aulas de língua portuguesa. Além disso, a escola pode criar espaços dedicados, como clubes de mídia e oficinas extracurriculares, onde os alunos possam praticar a verificação de informações e discutir temas atuais. Em comunidades mais vulneráveis, com menos acesso a tecnologias digitais, o foco deve estar no desenvolvimento de habilidades críticas que pode ser exercido por meios tradicionais de informação.

Outro ponto relevante é a disponibilidade de ferramentas automatizadas de verificação de informações, que podem ser facilmente acessadas de forma gratuita. Ferramentas como os *bots* de checagem analisados permitem que tanto professores quanto alunos tenham acesso a mecanismos que ajudam a identificar desinformação de maneira prática. A integração dessas ferramentas no processo educacional pode ser feita por meio de exercícios e projetos que incentivem o uso dessas tecnologias em pesquisas e verificações de notícias fraudulentas. Como são de acesso livre, essas ferramentas representam uma oportunidade para ampliar o alcance do combate à desinformação.

Além disso, a colaboração de escolas com universidades pode ajudar a direcionar essas ações de maneira mais objetiva, promovendo a criação de materiais didáticos e projetos de capacitação focados no letramento digital e na verificação de informações. Entre ferramentas automatizadas de impacto, é possível destacar algumas: o aplicativo GuaIA¹⁹, lançado pela Justiça Eleitoral em parceria com a Universidade Federal de Goiás (UFG) e com o Ministério Público de Goiás (MPGO), que detecta informações distorcidas e previne o uso indevido de tecnologias no contexto eleitoral; e o sistema uIAra²⁰, da Universidade Federal da Paraíba (UFPB), que identifica *deepfakes* de áudio.

¹⁹ O GuaIA foi lançado inicialmente para os integrantes do TRE-GO e para o Ministério Público Eleitoral. <https://diariodegoias.com.br/justica-eleitoral-em-parceria-com-a-ufg-lanca-aplicativo-para-combater-fake-news-nas-eleicoes/453098/>

²⁰ O uIAra tem o objetivo de auxiliar a justiça eleitoral e dar suporte às decisões dos magistrados no que diz respeito à desinformação (<https://www.ccs.ufpb.br/liev/lancamento-da-uiara-sistema-para-identificacao-de-deepfakes-de-audio-nas-eleicoes-de-2024/>)

5 CONSIDERAÇÕES FINAIS

Esta pesquisa explorou os desafios e potencialidades na detecção de notícias fraudulentas no contexto brasileiro, com foco na interação entre inteligência artificial e educação linguística. A crescente disseminação de desinformação nas plataformas digitais, como demonstrado ao longo do estudo, revela um problema multifacetado que transcende os aspectos tecnológicos e envolve também dimensões sociais, políticas e educacionais. Nesse cenário, a intersecção entre o uso de ferramentas de checagem de fatos automatizadas e o letramento digital nas escolas emergiu como um tema central para o enfrentamento da desinformação.

O estudo evidenciou que, embora as ferramentas automatizadas desempenhem um papel relevante na verificação de fatos, sua atuação enfrenta certas barreiras. O crescimento das agências de verificação e a incorporação de ferramentas de inteligência artificial em suas operações têm expandido o alcance dessas iniciativas. No entanto, conforme analisado, essas soluções ainda não são suficientes para lidar com a velocidade e o volume de desinformação circulante. A proliferação rápida de desinformação demanda um esforço contínuo de aperfeiçoamento das tecnologias de IA, como demonstrado pelas ferramentas analisadas ao longo do estudo, mas também de maior envolvimento social no combate à desinformação.

Em paralelo, o letramento digital nas escolas brasileiras mostrou-se um aspecto essencial para o combate à desinformação, não apenas como uma medida de prevenção a longo prazo, mas também como uma estratégia ativa de capacitação dos futuros cidadãos. O desenvolvimento de habilidades críticas e reflexivas para o consumo de informação digital é indispensável em uma era em que a produção e o compartilhamento de conteúdo são descentralizados, e a distinção entre o que é factual e o que não é torna-se cada vez mais complexa. A pesquisa sugere que o letramento digital deve ser incorporado de forma mais robusta e sistêmica nos currículos escolares, alinhado com as práticas de multiletramentos discutidas ao longo do trabalho. Tal iniciativa pode contribuir para a formação de estudantes capazes de identificar e resistir à manipulação informacional.

O uso de abordagens automatizadas, detalhado na análise das ferramentas de IA voltadas à detecção de desinformação, também se revelou promissor. As tecnologias baseadas em aprendizado de máquina, como o FakeCheck e o *bot* Fátima 3.0, podem complementar o trabalho humano de checagem, oferecendo escalabilidade e rapidez no processo de verificação. Entretanto, a pesquisa também ressalta que essas soluções automatizadas não são infalíveis e necessitam de constante aprimoramento, particularmente no que tange à interpretação de nuances culturais e linguísticas específicas do contexto brasileiro.

Assim, em consonância com Wardle e Derakhshan (2023), quaisquer soluções que visem enfrentar o distúrbio informacional de forma mais robusta e eficiente para o problema da desinformação precisam transcender ações isoladas e é necessário que estejam fundamentadas em uma abordagem multidisciplinar. A articulação entre verificação de fatos, letramento digital e o uso de IA configura-se como uma abordagem multidisciplinar e integrada para mitigar o bombardeio desinformativo. O combate à desinformação requer um esforço conjunto que envolve educação e tecnologia, bem como a conscientização crítica dos usuários das plataformas digitais. Em um país com alta penetração de *internet* e um cenário de desinformação política e social, a preparação de estudantes para lidar com essas questões torna-se um elemento crucial na construção de uma sociedade mais informada e resiliente. Por fim, este estudo contribui para o debate sobre desinformação ao sugerir que a interação entre tecnologia e educação pode oferecer caminhos viáveis para a contenção de notícias pré-fabricadas.

REFERÊNCIAS

ADORNO, T. W. **Minima moralia**: reflexões a partir da vida danificada. São Paulo: Ed. Ática, 1992.

AGÊNCIA LUPA. **Entenda nossas etiquetas**. São Paulo: Agência Lupa, 15 out. 2015. Disponível em: <https://lupa.uol.com.br/institucional/2015/10/15/entenda-nossas-etiquetas>. Acesso em: 18 fev. 2025.

ALLCOTT, Hunt; GENTZKOW, Matthew. **The state of applied econometrics**: Causality and policy evaluation. *Journal of Economic Perspectives*, v. 31, n. 2, p. 211–236, Spring 2017. DOI: 10.1257/jep.31.2.3. Acesso em: 07 set. 2024.

BENEVENTO, Fabrício. **Campanhas de desinformação no Brasil: lições aprendidas e desafios futuros - Fabrício Benevento**. YouTube, 03 mai. 2024. Disponível em: https://youtu.be/8WG_mTUEJbo. Acesso em 28 set. 2024.

BENTES, Ivana. Abralin em Cena: **Fake news e linguagem**. YouTube, 07 set. 2020. Disponível em: https://youtu.be/8WG_mTUEJbo. Acesso em: 28 set. 2024.

BOLTER, J. D. Writing space. Hillsdale, NJ: Erlbaum, 1991. _____. Hypertext and the question of visual literacy. In: REINKING, D. et al. (ed.). **Handbook of literacy and technology**: Transformations in a post-typographic world. Mahwah, NJ: Erlbaum, 1998. p. 3-13.

BUCCI, Eugênio. **Eugenio Bucci**: “É evidente que caminhamos para um jornalismo melhor”. Entrevista. Jornal O Povo, 7 jan. 2018. Disponível em: <https://mais.opovo.com.br/jornal/dom/2018/01/eugenio-bucci-e-evidente-que-caminhamos-para-um-jornalismo-melhor.html>. Acesso em: 26 ago. 2024.

BUZATO, Marcelo El Khouri. **Inteligência artificial, pós-humanismo e educação**: entre o simulacro e a assemblagem. *Dialogia*, São Paulo, n. 44, p. 1-11, jan./abr. 2023. DOI: 10.5585/44.2023.23906.

CAVALCANTE, Lidia Eugenia; SOUZA, Laiana Ferreira de. **Leitura, letramento digital e competência em informação**. *Revista Tecnologias na Educação*, ano 8, v. 17, p. 1-12, dez. 2016. Disponível em: <https://repositorio.ufc.br/handle/riufc/22737>. Acesso em: 22 set. 2024.

FARHANGIAN, A.; et al. **Fake news detection approaches**: challenges and opportunities. Journal of Machine Learning and AI, v. 10, n. 2, p. 45-60, 2024.

FERNANDES, C. M.; MONTUORI, C. **A rede de desinformação e a saúde em risco**: uma análise das fake news contidas em 'As 10 razões pelas quais você não deve vacinar seu filho'. Revista Eletrônica de Comunicação, Informação e Inovação em Saúde, v. 14, n. 2, 2020. DOI: 10.29397/reciis.v%25vi%25i.1975.

GARCIA, A. C. B. **Ética e inteligência artificial**. Computação Brasil, n. 43, p. 14–22, 2020. DOI: 10.5753/compbr.2020.43.1791.

GUO, Zhijiang; SCHLICHTKRULL, Michael; VLACHOS, Andreas. A Survey on Automated Fact-Checking. **Transactions of the Association for Computational Linguistics**, v. 10, p. 178-206, 2022. DOI: 10.1162/tacl_a_00454.

Fonte: LEITÃO, Matheus. **O golpe baixo e repugnante de Carla Zambelli contra Lula**. Veja, 25 abr. 2021. Disponível em: <https://veja.abril.com.br/coluna/matheus-leitao/o-golpe-baixo-e-repugnante-de-carla-zambelli-contra-lula>. Acesso em: 28 set. 2024.

LEMKE, Jay L. **Letramento metamidiático**: transformando significados e mídias. Trabalhos em Linguística Aplicada, Campinas, v. 49, n. 2, p. 455-479, jul./dez. 2010.

Lupa. **Entenda nossas etiquetas**. Lupa, 15 out. 2015. Disponível em: <https://lupa.uol.com.br/institucional/2015/10/15/entenda-nossas-etiquetas>. Acesso em: 28 set. 2024.

MACHADO, A. L.; FRANÇA, A. B. ; RANGEL, T. L. V. CARESTIA, MAPA DA FOME E O AGRAVAMENTO DA INSEGURANÇA ALIMENTAR E NUTRICIONAL EM TEMPOS DE PANDEMIA: O RETROCESSO BRASILEIRO NA POLÍTICA DE COMBATE À FOME. **Boletim de Conjuntura (BOCA)**, Boa Vista, v. 8, n. 24, p. 87–101, 2021. DOI: 10.5281/zenodo.5764610. Disponível em: <https://revista.ioles.com.br/boca/index.php/revista/article/view/512>. Acesso em: 9 out. 2024.

MACKINTOSH, Eliza. **Finland is winning the war on fake news**. What it's learned may be crucial to Western democracy. CNN, 2019. Disponível em:

<https://edition.cnn.com/interactive/2019/05/europe/finland-fake-news-intl/>. Acesso em: 08 set. 2024.

MAIA, Ari Fernando; URA, Carmen. **Fake news e a mobilização da paranoia:** colapso da individualidade e da democracia. **Dissonância:** revista de teoria crítica , v. 6, p. 378-420, 2022.

MARQUES, Igor; SALLES, Isadora; COUTO, João M. M.; PIMENTA, Breno C.; ASSIS, Samuel; REIS, Julio C. S.; SILVA, Ana Paula C.; ALMEIDA, Jussara M.; BENEVENUTO, Fabrício. **A Comprehensive Dataset of Brazilian Fact-Checking Stories.** *Journal of Information and Data Management*, v. 13, n. 1, p. 127-142, ago. 2022. DOI: 10.5753/jidm.2022.2354.

MINAYO, M. C. S. (org.). **Pesquisa social:** Teoria, método e criatividade. 18 ed. Petrópolis: Vozes, 2001.

MINISTÉRIO DA SAÚDE. **Formulário fake news.** Publicado em 13/07/2023. Disponível em: <https://www.gov.br/saude/pt-br/assuntos/saude-com-ciencia/formulario-fake-news>. Acesso em: 14 set. 2024.

MOLINA, M. D.; SUNDAR, S. S.; LE, T.; LEE, D. **“Fake News” Is Not Simply False Information: A Concept Explication and Taxonomy of Online Content.** *American Behavioral Scientist*, v. 65, n. 2, p. 180-212, 2021. DOI: 10.1177/0002764219878224. Acesso em: 18 fev. 2025.

MORENO, João; BRESSAN, Graça. FACTCK.BR: A New Dataset to Study Fake News. In: **Brazilian symposium on multimedia and the web**, 25., 2019, Rio de Janeiro. Anais [...]. New York: ACM, 2019. p. 1-3. DOI: 10.1145/3323503.3361698.

NASCIMENTO, F. R. S.; CAVALCANTI, G. D. C.; DA COSTA-ABREU, M. **Exploring Automatic Hate Speech Detection on Social Media:** A Focus on Content-Based Analysis.

Sage Open, v. 13, n. 2, 2023. DOI: 10.1177/21582440231181311.

NETO, J. J. S.; KARINO, C. A.; JESUS, G. R. de; ANDRADE, D. F. de. A infraestrutura das escolas públicas brasileiras de pequeno porte. *Revista do Serviço Público, /S. I.J*, v. 64, n. 3,

p. p. 377-391, 2014. DOI: 10.21874/rsp.v64i3.129. Disponível em: <https://revista.enap.gov.br/index.php/RSP/article/view/129>. Acesso em: 9 out. 2024.

NEWMAN, Nic; FLETCHER, Richard; KALOGEROPOULOS, Antonis; NIELSEN, Rasmus Kleis. **Reuters Institute Digital News Report 2019**. Reuters Institute for the Study of Journalism, 2019. Disponível em: https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2019-06/DNR_2019_FINAL_0.pdf. Acesso em: 28 ago. 2024.

O'NEIL, Cathy. **Weapons of math destruction**: how big data increases inequality and threatens democracy. New York: Crown Publishers, 2016. [livro eletrônico]

PIRES, Marília Freitas de Campos. **Multidisciplinaridade, interdisciplinaridade e transdisciplinaridade no ensino**. Interface - Comunicação, Saúde, Educação, Botucatu, v. 2, n. 2, p. 173-179, 1998. Disponível em: <https://www.scielosp.org/pdf/icse/v2n2/10.pdf>. Acesso em: 11 set. 2024.

REIS, Julio; MELO, Philipe; SILVA, Marcio; BENEVENUTO, Fabrício. **Desinformação em plataformas digitais: conceitos, abordagens tecnológicas e desafios**. In: Anais do XLII Congresso da Sociedade Brasileira de Computação. Porto Alegre: SBC, 2023. p. 1-24. DOI: 10.5753/sbc.12853.0.1.

REIS, Julio C. S.; MELO, Philipe; GARIMELLA, Kiran; BENEVENTO, Fabrício. **Can WhatsApp Benefit from Debunked Fact-Checked Stories to Reduce Misinformation?**. The Harvard Kennedy School (HKS) Misinformation Review, v. 1, n. 5, 2020. DOI: 10.37016/mr-2020-035. Acesso em: 24 set. 2024.

RIBEIRO, Filipe N.; SAHA, Koustuv; BABAEI, Mahmoudreza; HENRIQUE, Lucas; MESSIAS, Johnnatan; BENEVENUTO, Fabrício; GOGA, Oana; GUMMADI, Krishna P.; REDMILES, Elissa M. **On microtargeting socially divisive ads: A Case Study of Russia-Linked Ad Campaigns on Facebook**. In: FAT* '19: Conference on Fairness, Accountability, and Transparency, 29–31 January 2019, Atlanta, GA, USA. ACM, New York, NY, USA, 2019. DOI: 10.1145/3287560.3287580.

ROJO, Roxane. **Letramentos digitais**: a leitura como réplica ativa. Trabalhos em Linguística Aplicada, Campinas, v. 46, n. 1, p. 63-78, jan./jun. 2007.

ROJO, Roxane; MOURA, Eduardo (orgs.). Multiletramentos na escola. São Paulo: Parábola Editorial, 2012.

ROJO, Roxane; BARBOSA, Jacqueline P. **Hipermodernidade, multiletramentos e gêneros discursivos**. São Paulo: Parábola Editorial, 2015.

ROJO, R. Entre Plataformas, ODAs e Protótipos: **Novos multiletramentos em tempos de WEB2**. The ESPecialist, [S. l.], v. 38, n. 1, 2017. DOI: 10.23925/2318-7115.2017v38i1a2. Disponível em: <https://revistas.pucsp.br/index.php/esp/article/view/32219>. Acesso em: 8 out. 2024.

RUSSELL, S. J.; NORVIG, P. **Artificial intelligence**: a modern approach. 3 Ed. Upper Saddle River: Pearson Education Limited. 2010.

SANTAELLA, Lucia. **A inteligência artificial é inteligente?** São Paulo: Almedina, 2023. 178 p. ISBN 978-65-5427-054-0.

SEGURADO, Rosemary; CHICARINO, Tathiana Senne; CONCEIÇÃO, Desirée L. L. A percepção de conservadores e progressistas sobre memes desinformativos nas eleições 2020. Cadernos Metrópole, [S. l.], v. 24, n. 55, p. 1025–1049, 2022. DOI: 10.1590/2236-9996.2022-5508. Disponível em: <https://revistas.pucsp.br/index.php/metropole/article/view/56587>. Acesso em: 18 fev. 2025.

SCARINCI, Rui Gureghian; PALAZZO, José Moreira de Oliveira. **Extração de informação como base para descoberta de conhecimento em dados não estruturados**. Cadernos de Informática, v. 1, n. 1, p. 15-20, jan. 2001.

SEIBT, Taís. **Jornalismo de verificação como tipo ideal**: a prática de fact-checking no Brasil. 2019. 189 f. Tese (Doutorado em Comunicação) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2019. Disponível em: <https://lume.ufrgs.br/handle/10183/193359>. Acesso em: 30 ago. 2024.

SOUZA FILHO, L. A.; LAGE, D. A. **Entre ‘fake news’ e pós-verdade**: as controvérsias sobre vacinas na literatura científica. JCOMAL, v. 4, n. 2, 2021. DOI: 10.22323/3.04020901.

VLACHOS, Andreas; RIEDEL, Sebastian. Fact checking: Task definition and dataset construction. In: **Proceedings of the ACL 2014 Workshop on language technologies and computational social science**, p. 18–22, Baltimore, MD, USA, 2014. DOI: 10.3115/v1/W14-2508.

WARDLE, Claire; DERAKHSHAN, Hossein. **Desordem informacional**: para um quadro interdisciplinar de investigação e elaboração de políticas públicas. Tradução de Pedro Caetano Filho e Abilio Rodrigues. 1. ed. Campinas: Unicamp, Centro de Lógica, Epistemologia e História da Ciência, 2023. (Coleção CLE, v. 92). Disponível em: <https://www.cle.unicamp.br/ebooks/index.php/publicacoes/catalog/book/93>. Acesso em: 04 set. 2024.

ZUBOFF, S. “Big Other: capitalismo de vigilância e perspectivas para uma civilização de informação”. In: Bruno, F. (et al). **Tecnopolíticas da Vigilância**: perspectivas da margem. São Paulo: Boitempo, 2018.