

Caracterização da (Des)Informação das Narrativas sobre Vacinação em Grupos Públicos no WhatsApp no Brasil

Jônatas H. dos Santos¹, Julio C. S. Reis², Philipe de F. Melo², João F. H. Olivetti², Thales H. Silva¹,

Matheus Guimaraes¹, Glaucio de Souza¹, Marco A. G. Rodrigues¹,

Marcos A. Goncalves¹, Fabricio Benevenuto¹, Cristiano X. Lima¹

jonatashds@ufmg.br, jreis@ufv.br, philipe.freitas@ufv.br, joao.olivetti@ufv.br, thaleshenrique@ufmg.br

guimaraesmatheus22@gmail.com, glaucio.gss@gmail.com, magro.mg@terra.com.br

mgoncalv@dcc.ufmg.br, fabricio@dcc.ufmg.br, cxlima@ufmg.br

¹Universidade Federal de Minas Gerais (UFMG), Brasil

²Universidade Federal de Viçosa (UFV), Brasil

ABSTRACT

WhatsApp has become one of the main channels for disseminating health information in Brazil, especially during the COVID-19 pandemic. In this work, we investigate the spread of (mis)information about vaccines in public groups on the platform, focusing on identifying dominant narratives, engagement patterns, and connections to other online information sources. Using a set of 84,640 messages collected between 2020 and 2023, we applied Natural Language Processing (NLP) techniques, including topic modeling (BERTopic) and textual similarity analysis. Overall, the results show that alarmist, misinformative, and politicized content is prevalent, with many messages repeated across multiple groups. We also observed a concentration of activity among a few highly active users, suggesting coordinated amplification strategies. This analysis helps to understand the mechanisms of misinformation spread in private messaging environments and provides insights for public policy and digital health countermeasures.

KEYWORDS

WhatsApp, Saúde, Vacina, COVID-19, Desinformação, *Fake News*

1 INTRODUÇÃO

O Brasil enfrenta, em 2025, um novo alerta sanitário dado que o país retornou à lista das 20 nações com o maior número de crianças não vacinadas no mundo, segundo relatório conjunto da UNICEF e da Organização Mundial da Saúde (OMS) [1]. A diminuição das taxas de vacinação no país nestes últimos anos é preocupante, considerando o histórico do país em campanhas de imunização em massa. Entre um dos fatores atuais que contribuem em peso para esse cenário, destaca-se o papel da desinformação sobre saúde [2], que reduz a confiança da população em vacinas e nas instituições competentes.

Em meio a esse desafio, o WhatsApp se consolida como um canal central deste ambiente (des)informacional brasileiro [3]. Com mais de 120 milhões de usuários no país e altamente difundido em todas as regiões do Brasil e entre diferentes perfis demográficos, a plataforma é um ambiente fértil para a propagação de campanhas de desinformação, especialmente devido à sua arquitetura mais

fechada e restrita [4]. O formato de comunicação mais privada e criptografada dos grupos (em comparação a outras redes como Facebook e Twitter/X) amplia os desafios de monitoramento, dificultando a identificação de padrões e estratégias de manipulação do conteúdo compartilhado pelo aplicativo e colocando o WhatsApp no centro do debate sobre *fake news* no país [5].

Embora a plataforma seja amplamente apontada como vetor de discussão política [6–8], a operacionalização da plataforma sobre questões relacionadas à saúde também ganhou força durante a pandemia de COVID-19 [9–11], quando o debate sobre vacinas tornou-se um dos principais focos deste espaço digital de mensagens [12], mobilizando discursos conspiratórios, ataques à ciência e teorias sobre supostos riscos à saúde, trazendo alarmismo e medo à população. A estrutura dos grupos públicos no WhatsApp, acessíveis por convites amplamente compartilhados na Web, permitiu a formação de grandes comunidades dedicadas à disseminação desses tipos de conteúdos, expandindo-se com força e alcance em diferentes segmentos sociais, mas sem uma resposta institucional proporcional por parte das autoridades [13].

Desta forma, compreender como essas mensagens circulam dentro do WhatsApp, quem as envia, quais temas predominam, as narrativas criadas e de onde vêm essas informações é essencial para o enfrentamento da crise vacinal. Neste contexto, a presente pesquisa busca contribuir com esse esforço ao avançar no entendimento da circulação de (des)informação sobre vacinas no ambiente de comunicação fechada e de difícil acesso do WhatsApp e contribuir para o desenvolvimento de estratégias de enfrentamento mais eficazes, tanto do ponto de vista técnico quanto de políticas públicas. Mais especificamente, este trabalho tem como objetivo responder às seguintes questões de pesquisa (QPs):

- **QP1:** Como o volume e os padrões de engajamento (por usuário e por grupo) das discussões sobre vacina evoluíram ao longo do tempo?
- **QP2:** Quais foram as principais narrativas e tópicos de discussão sobre vacinas nos grupos públicos de WhatsApp no Brasil durante a pandemia de COVID-19?
- **QP3:** Como os links sobre vacina no WhatsApp se conectam a outras fontes de dados da Web e plataformas sociais?

Para responder às questões de pesquisa propostas, adotou-se uma abordagem metodológica quantitativa fundamentada em técnicas de Processamento de Linguagem Natural (PLN), aplicada a uma base de dados real e de grande escala composta por 84.640

In: Proceedings of the Brazilian Symposium on Multimedia and the Web (WebMedia'2025). Rio de Janeiro, Brazil. Porto Alegre: Brazilian Computer Society, 2025.

© 2025 SBC – Brazilian Computing Society.

ISSN 2966-2753

mensagens relacionadas à temática da vacinação, coletadas em grupos públicos do WhatsApp no período de 2020 a 2023, submetidas a um processo de filtragem e pré-processamento, visando à consolidação de um conjunto de dados estruturado e coerente. Em seguida, conduziu-se uma análise descritiva das características temporais e dos padrões geográficos de envio das mensagens, com ênfase tanto na atividade dos usuários quanto na dinâmica de circulação da informação entre os grupos. Posteriormente, procedeu-se a uma análise mais aprofundada do conteúdo textual, a partir da identificação de termos mais frequentes e da modelagem de tópicos para a extração e agrupamento temático, o que permitiu a identificação de tópicos recorrentes e de narrativas predominantes ao longo do período investigado. Por fim, analisaram-se também os links incorporados nas mensagens, com o objetivo de mapear conexões entre os conteúdos compartilhados no WhatsApp e outras esferas do ecossistema informacional.

Os resultados indicam esforços deliberados de amplificação de conteúdos desinformativos no WhatsApp, com destaque para uma desigualdade geográfica: embora grupos de todas as regiões tenham sido identificados, a maior atividade ocorreu no Sudeste e Nordeste. Também se observou participação significativa de números estrangeiros, sugerindo possíveis estratégias transnacionais. As mensagens analisadas concentraram-se em narrativas que questionam a eficácia das vacinas, promovem tratamentos sem base científica e estimulam desconfiança nas autoridades sanitárias. Os links compartilhados provinham majoritariamente de fontes de baixa credibilidade, como sites alternativos e blogs, reforçando o papel do WhatsApp como vetor de entrada e amplificação de conteúdos desinformativos. Em conjunto, os achados evidenciam a plataforma como infraestrutura estratégica na circulação de (des)informação em saúde e revelam padrões de comportamento informacional que podem subsidiar o desenvolvimento de políticas públicas e estratégias de intervenção mais eficazes.

O restante deste artigo está organizado da seguinte forma. A Seção 2 apresenta um panorama da literatura sobre desinformação no WhatsApp, em especial sobre saúde e as lacunas de pesquisa presentes na área. A Seção 3 detalha a metodologia de coleta, filtragem e análise dos dados do WhatsApp. Em seguida, os resultados são apresentados e discutidos na Seção 4. Por fim, a Seção 5 sumariza as conclusões do estudo e aponta direções para trabalhos futuros.

2 TRABALHOS RELACIONADOS

A hesitação vacinal, um desafio histórico de saúde pública, foi intensificada pelo ambiente virtual contemporâneo. Antes mesmo da pandemia de COVID-19, a OMS já a classificava como uma das principais ameaças à saúde global [14]. Com a expansão das redes sociais, instaurou-se a chamada “infodemia” [15, 16], marcada pela disseminação massiva de informações falsas com efeitos negativos sobre a saúde pública. Nesse contexto, esta seção revisa a literatura a partir de duas abordagens: (i) a anatomia das narrativas antivacinas e (ii) o ecossistema digital de desinformação no WhatsApp.

2.1 Anatomia das Narrativas Antivacinas

A pesquisa acadêmica tem identificado um conjunto de pilares temáticos que formam o núcleo do movimento antivacina online

atual. Esses temas são adaptáveis, mas mantêm uma estrutura fundamental em diferentes plataformas e contextos. Um dos pilares centrais é a desinformação médica e de segurança, que consiste em fabricar ou exagerar os riscos associados às vacinas. Por exemplo, alegações falsas sobre um suposto elo entre a vacina Tríplice Viral (MMR) e o autismo continuam sendo um argumento utilizado [17]. Além disso, com o surgimento das vacinas de COVID-19, houve alegações de que tais vacinas causariam infertilidade, alterações nocivas ao DNA ou levariam a mortes súbitas [17].

Intimamente ligadas a tais preocupações estão as teorias da conspiração. Essas narrativas atribuem intenções maliciosas a atores poderosos. Isso inclui a ideia de que as vacinas seriam uma ferramenta de controle governamental através da implantação de microchips para vigilância em massa, ou que fazem parte de um plano secreto da elite global para consolidar poder [18]. Outra teoria é que a indústria farmacêutica estaria ocultando os riscos das vacinas para maximizar seus lucros [19]. Outro pilar ataca a eficácia e o desenvolvimento das vacinas, argumentando que o processo de produção teria ocorrido de maneira “apressada” e sem os devidos testes. Finalmente, um pilar altamente politizado é o apelo à liberdade individual e à autonomia corporal. Os incentivos à vacinação são vistos como uma forma de autoritarismo governamental e violação dos direitos individuais [18].

A eficácia das estratégias desinformativas identificadas reside não apenas no conteúdo veiculado, mas também na forma como as mensagens são estruturadas e comunicadas. Um recurso recorrente envolve a menção a supostos médicos e cientistas, com o intuito de conferir uma aparência de autoridade e credibilidade às alegações apresentadas [20]. Outra estratégia de alto impacto consiste na mobilização de apelos emocionais, especialmente por meio de relatos pessoais que descrevem, em detalhes, experiências negativas atribuídas à vacinação. Tais narrativas exploram mecanismos psicológicos de persuasão, evidenciando que o engajamento emocional pode aumentar significativamente a efetividade da desinformação [21].

Além das narrativas disseminadas, fatores psicológicos influenciam o compartilhamento de desinformação. Destacam-se vieses cognitivos como o de confirmação e o raciocínio motivado [22], além de baixos níveis de reflexão e mentalidade conspiratória, associados à maior resistência à vacinação [23]. Emoções como paranoia e ansiedade diante da morte também exercem influência [24]. A confiança surge como preditor central da adesão vacinal, sendo a hesitação ligada à desconfiança em instituições científicas e governamentais [25], especialmente em grupos socialmente vulneráveis [26]. Ademais, a vacinação tem sido cada vez mais associada à identidade política, refletindo os efeitos da polarização partidária [27].

O movimento antivacina e, de modo mais amplo, a desinformação sobre saúde extrapolam o ambiente digital, produzindo impactos concretos no mundo real. Evidências apontam sua associação a quedas nas taxas de vacinação [2]. Um estudo, por exemplo, estimou que um aumento de dez pontos percentuais no sentimento anti-vacina online correspondeu a uma redução de 0,43% na cobertura da vacina MMR em municípios italianos. Além de afetar diretamente a incidência de doenças evitáveis, o ceticismo vacinal contribui para o aprofundamento de divisões sociais e para a erosão da confiança nas instituições [28].

2.2 Ecossistema de Desinformação do WhatsApp

Nesse contexto, as plataformas sociais não atuam como meros canais neutros de disseminação de informação. Ao contrário, sua arquitetura técnica e modelo de negócios frequentemente contribuem para a amplificação da desinformação [29]. A lógica algorítmica predominante, baseada na maximização do engajamento dos usuários, tende a priorizar conteúdos que geram reações intensas, como curtidas, comentários e compartilhamentos. Dado que a desinformação frequentemente adota uma linguagem sensacionalista, com forte apelo emocional, ela apresenta maior probabilidade de ser promovida por esses mecanismos.

A disseminação de desinformação também pode ser deliberadamente promovida por agentes maliciosos. Contas automatizadas (*bots*) e perfis inautênticos (*trolls*) têm sido associados à manipulação do debate público sobre vacinas [30]. Além disso, a formação de câmaras de eco nas redes sociais expõe os usuários a conteúdos que reforçam suas crenças prévias, intensificando o viés de confirmação e a polarização ideológica [31].

Dentre essas plataformas, o WhatsApp ocupa uma posição singular no ecossistema comunicacional brasileiro, em virtude de sua ampla penetração e popularidade no país. A aplicação é utilizada não apenas para interações interpessoais, mas também como meio de comunicação em grandes grupos públicos, nos quais circulam conteúdos que vão desde elementos de entretenimento, como memes e figurinhas, até discussões sobre temas sensíveis, como saúde e política [32]. Essa evolução configura uma transição do uso original da plataforma – centrado em mensagens privadas entre indivíduos – para sua função como canal de comunicação pública e de alto potencial viral [4].

Todavia, a arquitetura do WhatsApp, caracterizada por comunicação criptografada de ponta a ponta, descentralização e ausência de indexação pública, impõe severas limitações ao monitoramento automatizado e em larga escala [33]. Em contraste com plataformas sociais de natureza mais aberta, como Twitter (X) ou Facebook, o WhatsApp apresenta barreiras significativas à rastreabilidade de conteúdo e à auditoria pública. Esse contexto favorece a circulação de informações não verificadas, incluindo boatos, teorias conspiratórias e campanhas coordenadas de desinformação [4, 34–36].

Diversos estudos apontam o WhatsApp como um dos principais vetores de campanhas de desinformação em contextos políticos, especialmente nas eleições brasileiras de 2018 e 2022 [6, 8, 37–39]. Nesses períodos, conteúdos enganosos circularam amplamente em grupos públicos e privados [3, 4], revelando um novo paradigma de influência política baseado em estratégias organizadas, como o uso de listas de transmissão, o compartilhamento massivo e a articulação entre múltiplos grupos interconectados [34]. O reenvio sucessivo de mensagens ampliou ainda mais o alcance desses conteúdos, produzindo cascatas informacionais que atravessaram grupos e regiões [34, 40].

Durante a pandemia de COVID-19, o WhatsApp reafirmou sua centralidade como uma das principais vias de disseminação de desinformação sobre saúde, tanto no Brasil [9] quanto em outros contextos internacionais [41–43], com particular ênfase em temas relacionados à vacinação [44]. Nesse cenário, a circulação de conteúdos multimodais – como imagens, vídeos e áudios – com estética

informal e forte apelo emocional tornou-se recorrente, dificultando o trabalho de checagem por agências de verificação de fatos e pelas próprias plataformas que atuam no combate à desinformação [45].

Sob o ponto de vista metodológico, a extração e análise de dados do WhatsApp ainda enfrentam desafios técnicos. A maioria dos estudos adota abordagens qualitativas, como entrevistas [46, 47] ou utiliza amostras limitadas de usuários [41, 48–50]. Para superar essas limitações, surgiram estratégias de coleta em larga escala, como o ingresso automatizado em grupos públicos [51], além do uso de sistemas de monitoramento dedicados [52, 53], que fortalecem o aparato metodológico e viabilizam análises quantitativas mais abrangentes.

2.3 Lacuna de Pesquisa

Apesar desses avanços, ainda são limitadas as investigações que relacionam a desinformação no WhatsApp ao contexto específico da vacinação, em especial de maneira longitudinal e geograficamente contextualizada. Muitos estudos se concentram em eventos ou campanhas pontuais, sem explorar a evolução do discurso ao longo do tempo e sua relação com fatores como território, usuários-chave e ambientes externos (como websites e outras plataformas). Neste sentido, nosso trabalho avança o estado da arte ao apresentar uma análise em larga escala de mensagens sobre vacinas em grupos públicos de WhatsApp durante quatro anos, explorando aspectos temporais, geográficos, discursivos e hipertextuais de maneira integrada.

3 METODOLOGIA

Nesta seção, são detalhados os procedimentos metodológicos adotados neste estudo, desde a construção e filtragem do conjunto de dados até as etapas de análise exploratória e semântica das mensagens. Conforme ilustrado na Figura 1, o *pipeline* compreende a coleta automatizada de mensagens em grupos públicos de WhatsApp, a filtragem para selecionar conteúdos relacionados à vacinação e a consolidação da base de dados. Em seguida, descreve-se também as análises realizadas, incluindo a investigação temporal e geográfica das mensagens, a modelagem de tópicos e o estudo das URLs compartilhadas. Por fim, são destacadas as principais limitações associadas aos dados.

3.1 Monitoramento e Coleta de Dados

A base de dados utilizada neste trabalho foi obtida a partir de trabalhos disponibilizados por autores de [7, 54], responsáveis pela construção de um grande repositório de mensagens enviadas em grupos públicos do WhatsApp. A coleta foi realizada entre 2020 e 2023 e abrange mensagens de grupos públicos acessíveis via links de convite do tipo `chat.whatsapp.com/{ID}`. Tais links foram obtidos por meio de busca ativa na Web e plataformas sociais, utilizando um dicionário de palavras-chave relacionadas ao contexto geopolítico brasileiro, e.g., nomes de todos os políticos, partidos, bem como palavras associadas ao extremismo. Para realizar a extração dos dados, foi empregada uma abordagem automatizada e em larga escala, conectada a números de telefone e usuários reais da plataforma. Neste processo, são salvas as mensagens recebidas e registrados os metadados como ID da mensagem, autor, grupo, data/hora, tipo de mídia e conteúdo textual, passando por processos éticos para anonimizar os dados segundo as normas da Lei Geral de Proteção de Dados Pessoais (LGPD). Este processo de anonimização é uma etapa

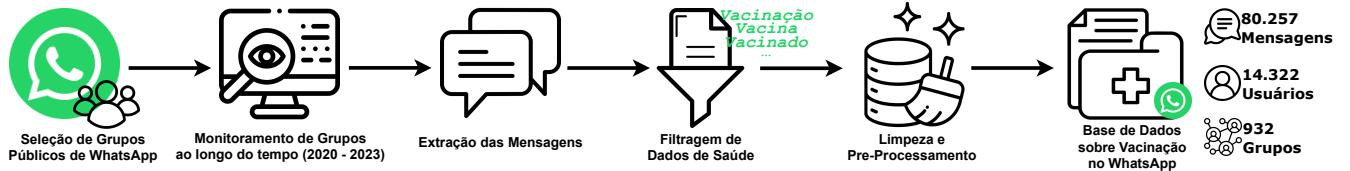


Figura 1: Diagrama geral do *pipeline* de coleta e filtragem dos dados sobre vacinação do WhatsApp.

Tabela 1: Palavras-chave do filtro sobre vacina.

Palavra-chave	#Ocorrências	Usuários únicos	Grupos únicos
vacina	65373	10710	876
vacinas	32476	6835	768
vacinação	17773	4450	695
vacinados	7813	2521	533
vacinar	5675	2458	532
vacinado	3504	1758	473
vacinadas	2517	1058	365
vacinada	1928	990	348
vacinal	1686	855	340
vacinou	1637	1012	335
vachina	1487	670	274
vacinei	147	130	94
v4c1n4	14	8	8

essencial em pesquisas envolvendo plataformas de mensagens instantâneas que garante segurança dos dados sensíveis dos usuários do grupo, enquanto também preserva informações relevantes como o código de área (DDD) e o código internacional (DDI) do remetente, permitindo análise geográfica sem violar a privacidade. As mensagens do repositório foram então recuperadas em formato estruturado (JSON) e armazenadas para análise.

3.2 Filtragem de Dados de Saúde

A base bruta considerando o período analisado (i.e., 2020–2023) é composta de milhões de mensagens, sendo a grande maioria delas não relacionadas ao foco sobre o tema de saúde desta pesquisa. Para delimitar o escopo às discussões sobre vacinação, aplicou-se um processo de filtragem baseado em palavras-chave. Dada a variedade dos dados e grande quantidade de ruídos na base, nossa filtragem considerou termos bem diretos (e.g., “vacina”, “vacinação”, “vacinado”) e também grafias alternativas e gírias populares, como “vachina” e “v4c1n4”, comumente associadas a discursos conspiratórios ou negativos. Empiricamente, outros termos associados (como os nomes das vacinas e outros termos médicos) foram testados na busca, porém foram removidos após os resultados destes retornarem também conteúdos fora do escopo. O resultado desta filtragem foi um subconjunto contendo 84.640 mensagens, postadas por 15.148 usuários. A Tabela 1 apresenta as palavras-chave utilizadas e a distribuição de mensagens por termo. Os termos alternativos como “vachina” (com 1.487 ocorrências) e “v4c1n4” (com 14 ocorrências) revelam uma dimensão discursiva pejorativa e criptografada da desinformação, frequentemente associada a posicionamentos ideológicos críticos às vacinas, especialmente àquelas da China.

Embora o conteúdo coletado seja de natureza predominantemente política, grandes grupos públicos de WhatsApp foram veotores de vários outros tópicos de relevância no país e constituíram

uma das principais formas de comunicação no país no período analisado [3]. Apesar disso, mesmo sendo um dos maiores dataset de WhatsApp sobre o tópico de vacina que se tem conhecimento, devido à ausência de bases de dados oficiais ou mecanismos de acesso estruturado ao conteúdo do WhatsApp, não é possível aferir com precisão a representatividade estatística da amostra analisada.

3.3 Limpeza e Pré-processamento dos Dados

O conjunto de dados bruto utilizado não fornece dados agregados de compartilhamento sobre as mensagens nem de cópias. Para garantir a qualidade das informações extraídas a partir da base e evitar redundâncias, foi aplicada uma etapa de remoção de duplicatas baseada em similaridade semântica. Foi utilizada a biblioteca SentenceTransformers [55] para gerar *embeddings* vetoriais das mensagens e a biblioteca Faiss [56] para indexação e busca por vizinhos mais próximos, com base em distância euclidiana.

Empiricamente, foi identificado que a maior concentração de mensagens duplicadas estava em distâncias inferiores a 0,02. Assim, adotou-se esse valor como limiar para considerar mensagens como duplicadas. Para cada grupo de mensagens semelhantes, apenas a primeira ocorrência foi mantida. Nesta etapa foram descartadas 5,18% das mensagens do conjunto de dados bruto por se tratarem de duplicadas. Ao final, foi construído um conjunto de dados contendo 80.257 mensagens únicas, postadas por 14.322 usuários em 932 grupos públicos distintos no WhatsApp.

3.4 Análise Temporal e Geográfica dos Dados

Depois de consolidar o conjunto de dados, partiu-se para algumas análises dos dados. O primeiro passo neste sentido foi uma análise temporal que considerou a distribuição das mensagens ao longo dos anos de coleta, permitindo observar alterações de volume e possíveis correlações com eventos sociopolíticos ou sanitários. Quanto à origem geográfica das mensagens, foram explorados os códigos DDD (discagem direta à distância) e DDI (discagem direta internacional) dos números de telefone dos usuários, preservados no processo de anonimização. A partir desses códigos, foi possível mapear a distribuição das mensagens por estado brasileiro e por país.

3.5 Modelagem de Tópicos

Para compreender melhor as principais narrativas em torno de vacinação no WhatsApp presentes nas mensagens, foi realizada uma análise exploratória do conteúdo textual a partir de técnicas de Processamento de Linguagem Natural (PLN) para avaliar de forma quantitativa o conteúdo compartilhado e obter uma visão preliminar dos temas em circulação. Para isso, a frequência dos termos mais recorrentes foi computada no conjunto de dados e gerou-se uma nuvem de palavras com as expressões mais frequentes.

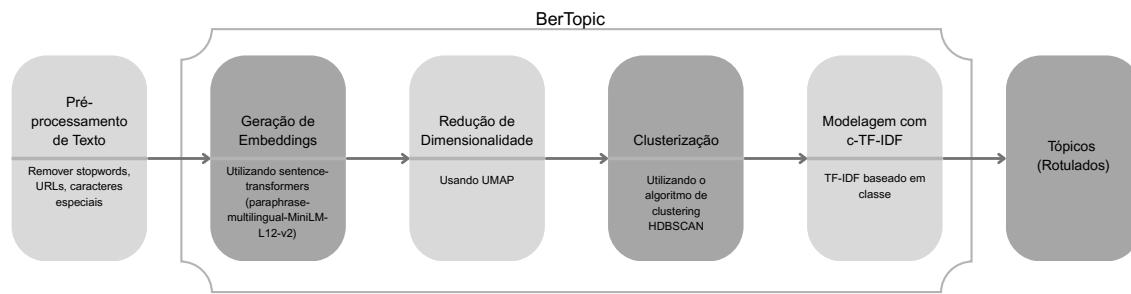


Figura 2: Fluxograma das etapas internas do BERTopic.

Em seguida, com o objetivo de aprofundar essa análise temática, foi aplicada uma técnica de modelagem de tópicos baseada no algoritmo *BERTopic* [57]. O *BERTopic* é uma abordagem que combina representações semânticas (*embeddings*) geradas por modelos *Transformers* com técnicas de redução de dimensionalidade, clusterização e representação de tópicos baseada em frequência de termos.

Os passos desta modelagem estão ilustrados na Figura 2 e seguem as seguintes etapas: **Pré-processamento textual:** As mensagens foram inicialmente tratadas com a remoção de elementos que pudessem interferir na análise. **Geração de embeddings:** Foi utilizado o modelo *paraphrase-multilingual-MiniLM-L12-v2* [58], da biblioteca *sentence-transformers*, para converter as mensagens em vetores de alta dimensionalidade. Esse modelo é otimizado para capturar similaridade semântica em sentenças curtas, sendo apropriado para o contexto mais informal das mensagens de WhatsApp. **Redução de dimensionalidade:** Os vetores gerados foram então reduzidos em dimensionalidade por meio do algoritmo *Uniform Manifold Approximation and Projection* (UMAP) [59], com o objetivo de preservar a estrutura local e global dos dados. **Clusterização:** Foi aplicado o algoritmo *Hierarchical Density-Based Spatial Clustering of Applications with Noise* (HDBSCAN) [60] sobre os *embeddings* reduzidos para identificar agrupamentos de mensagens semanticamente similares. **Modelagem c-TF-IDF:** Em seguida, foi usado o método *class-based TF-IDF* para representar os tópicos. Essa técnica calcula a importância de cada termo dentro do conjunto de documentos que compõem cada cluster, permitindo extrair palavras-chave mais representativas de forma precisa. Isso resulta na construção de tópicos semanticamente coesos e de fácil interpretação. **Rotulação:** Por fim, foi realizada a rotulagem de cada um dos tópicos com o auxílio da LLM em código aberto GEMMA3 12B [61] da Google, com base nas palavras mais frequentes de cada um.

3.6 Extração de URLs

Para investigar a circulação de conteúdos externos nas mensagens analisadas, foram extraídas todas as URLs presentes nas mensagens. A análise dessas URLs visa entender como os grupos públicos de WhatsApp se conectam a outros ambientes informacionais na Web, revelando possíveis padrões e relacionamentos entre plataformas, um fenômeno conhecido como *cross-platform dissemination* [62]. Para isso, foi utilizada a biblioteca *BeautifulSoup*¹ para realizar a expansão e normalização das URLs, permitindo a identificação precisa do domínio principal de cada link. A partir desses dados,

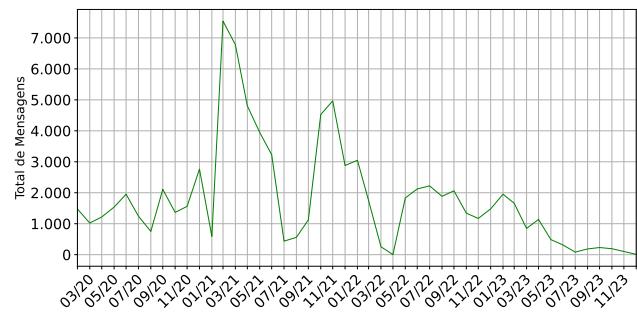


Figura 3: Volume total de mensagens enviadas (2021–2023).

foi construída uma lista com os domínios mais frequentemente compartilhados nas mensagens.

3.7 Limitações

Existem limitações relacionadas ao processo de investigação. A coleta de dados foi restrita a grupos públicos de WhatsApp acessíveis por convites indexados, o que pode não capturar a totalidade das interações privadas ou de comunidades menores. Além disso, a análise temporal abrange o período de 2020 a 2023, limitando a observação de mudanças posteriores na dinâmica das narrativas e na plataforma.

4 RESULTADOS

Nesta seção, são apresentadas as principais descobertas obtidas a partir da análise quantitativa e qualitativa das mais de 80 mil mensagens sobre vacinação compartilhadas em grupos públicos de WhatsApp entre 2020 e 2023 em relação a diferentes dimensões do fenômeno observado: como a evolução temporal do debate (Seção 4.1), origem e/ou distribuição geográfica (Seção 4.2), tópicos dominantes identificados via modelagem semântica (Seção 4.3), e finalmente, fontes externas compartilhadas via URLs (Seção 4.4).

4.1 Evolução Temporal dos Dados

Inicialmente, investigou-se a distribuição temporal das mensagens relacionadas à vacinação ao longo do período analisado. A Figura 3 mostra a quantidade mensal de mensagens entre janeiro de 2020 e dezembro de 2023, com destaque a dois picos de atividade significativos, ambos no ano de 2021. O primeiro e mais expressivo pico ocorre por volta de março de 2021, com mais de 7.000 mensagens publicadas. Esse aumento repentino coincide com a chegada das

¹Disponível em: <https://www.crummy.com/software/BeautifulSoup/bs4/doc>

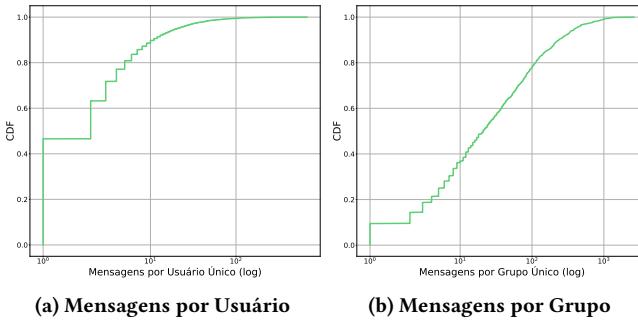


Figura 4: CDF do volume de mensagens enviadas.

primeiras doses da vacina contra COVID-19 no Brasil [63] e a aplicação da primeira dose em território nacional na enfermeira Mônica Calazans [64], um evento que teve grande repercussão na mídia nacional e simbólica para o início da campanha de imunização no país. Esse pico pode ser interpretado como um marco discursivo também dentro do WhatsApp no país, em que diversas narrativas (tanto de apoio, hesitação ou rejeição) começaram a circular intensamente nos grupos observados.

Já o segundo pico relevante é observado ao final de 2021, com aproximadamente 5.000 mensagens. Este momento coincide com dois fatores importantes: uma nova onda da pandemia com o auge do número de mortes causadas pelo impacto da variante Ômicron no Brasil [65] e a aprovação, pela ANVISA, do uso emergencial da vacina da Pfizer para crianças entre 5 e 11 anos [66]. Foi observado que essa introdução da vacinação pediátrica, gerou novos ciclos de debate no WhatsApp, marcados por um volume considerável de desinformação e discursos de pânico moral em torno da segurança da vacina para crianças [67], além da amplificação de teorias conspiratórias. Após o final de 2022, observa-se uma tendência de queda progressiva no volume de mensagens relacionadas à vacinação, o que pode indicar um esgotamento da pauta nos grupos analisados, ou a transição da desinformação para outros temas políticos e sociais dominantes no período.

4.2 Perfil e Origem das Mensagens

Em seguida, analisou-se a atividade entre usuários e grupos, bem como a origem geográfica das mensagens. A Figura 4 apresenta a Função de Distribuição Acumulada (CDF) da quantidade de mensagens por usuário e por grupo. Pode-se observar um padrão altamente desigual de participação entre os usuários, no qual cerca de 45% dos usuários enviaram apenas uma única mensagem relacionada à vacinação, e aproximadamente 90% contribuíram com no máximo 10 mensagens. Esse comportamento sugere que a maioria dos usuários interage pontualmente com o tema. Por outro lado, uma fração dos usuários demonstrou uma participação significativamente mais intensa. Há indivíduos que, cada um, isoladamente, publicou centenas de mensagens, evidenciando uma certa concentração da produção e circulação de conteúdo na mão de poucos autores na plataforma, o que potencialmente indica uma tentativa de liberada de amplificação do discurso ou de campanhas coordenadas.

Quanto à distribuição por grupo, os dados revelam uma dinâmica um pouco menos desigual: cerca de 10% dos grupos continham apenas uma mensagem, enquanto aproximadamente 80% apresentaram

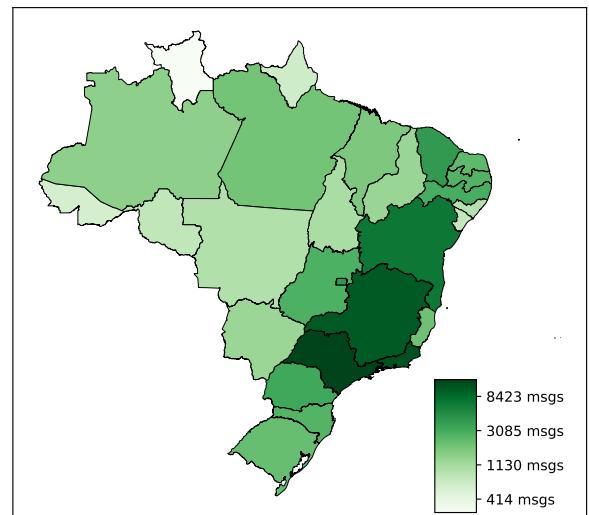


Figura 5: Mapa coroplético com volume de mensagens/estado.

até 100 mensagens. Essa distribuição mais fragmentada sugere que, apesar de existirem grupos com alta atividade, o debate sobre vacinação esteve mais presente e de maneira relativamente mais espalhada no conjunto de dados.

Com relação à origem geográfica das mensagens, foi notada uma ampla cobertura nacional, com participação de usuários de todos os estados brasileiros. Os dados mostram que os estados com maior volume de mensagens foram São Paulo (SP) (13.190 mensagens), Rio de Janeiro (RJ) (11.282) e Minas Gerais (MG) (10.180), seguidos por Bahia (BA), Distrito Federal (DF) e Ceará (CE). Essa concentração acompanha, em parte, a distribuição populacional do país, mas também reflete o papel central dessas regiões nos debates políticos durante a pandemia, com destaque para o DF, que apesar de ser apenas a vigésima unidade federativa mais populosa, configura-se como o 5º com mais mensagens. A Figura 5 apresenta um mapa coroplético que evidencia a dispersão geográfica das mensagens. Nota-se que, mesmo em estados com menor densidade populacional, como Acre (AC), Roraima (RR) e Amapá (AM), houve participação relevante.

Além da distribuição nacional, foi detectado um volume substancial de mensagens originadas de números internacionais. Uma análise dos DDIs mostra que países como Estados Unidos (463 mensagens), Portugal (209) e Alemanha (90) figuram entre as principais origens estrangeiras. Ao todo, foram identificadas mensagens provenientes de mais de 20 países, o que representa um fluxo significativo de informação transnacional. Embora boa parte pareça ser de usuários que moram no exterior, esse volume também pode indicar a presença de contas não legítimas influenciando na discussão.

4.3 Narrativas sobre Vacinação no WhatsApp

Antes de aplicar a modelagem semântica por tópicos, foi realizada uma análise exploratória inicial baseada na frequência de termos presentes nas mensagens sobre vacinação. A Figura 6 apresenta uma nuvem de palavras com os termos mais frequentes após a remoção de stopwords, links, símbolos e pontuações. Os dados sugerem a existência de um ambiente discursivo altamente polarizado, em que elementos ligados à saúde pública (“vacina”, “saúde”, “ministério”, “contra”, “Covid”) aparecem lado a lado com entidades políticas (“Jair Bolsonaro”, “polícia federal”, “Estados Unidos”, “China”). Para além

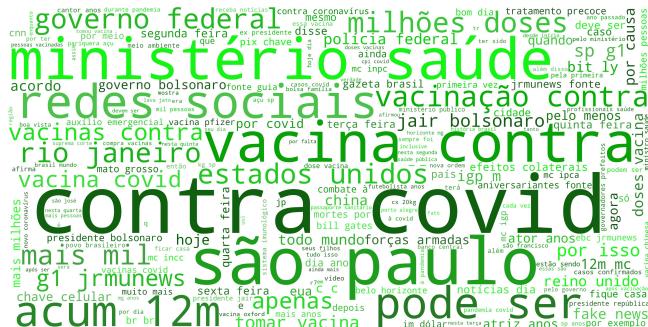


Figura 6: Nuvem de palavras contendo os termos mais populares no dataset após filtragem de *stopwords* e links.

disso, observa-se entre termos recorrentes plataformas digitais (“redes sociais”, “Facebook”, “YouTube”) e em menor escala, conceitos genéricos de grande carga ideológica (“tratamento precoce”, “fique casa”, “liberdade”, “direito”, “povo”). Além disso, a menção recorrente ao então presidente Jair Bolsonaro sugere uma centralidade da figura política nos debates sobre vacinação, tanto na mobilização de seus apoiadores contrários à obrigatoriedade da vacina, como também aos que criticam sua postura frente à pandemia.

Os tópicos extraídos pelo BERTopic, apresentados na Tabela 2, revelam uma série de padrões preocupantes, em que vários refletem narrativas alarmistas ou potencialmente desinformativas, como a

associação entre mortes e vacinas, desconfiança na vacina chinesa e supostos efeitos adversos (e.g., Tópico #01). Também foram identificados conteúdos sobre vacinação infantil, usuários que buscavam atacar a vacinação contra a gripe ou duvidavam do número de doses aplicadas até aquele momento. Um exemplo de mensagem para cada um dos tópicos pode ser visto na Tabela 3, essas mensagens geralmente são acompanhadas de dúvidas ou preocupações exageradas, em busca de incitar o medo aos leitores. Além disso, notou-se intensa circulação de vídeos e links de canais de notícias.

Essa categorização evidencia o papel do WhatsApp na difusão de mensagens que podem afetar negativamente a confiança pública nas vacinas. Observa-se também a presença de conteúdos que podem articular desinformação com aspectos políticos e institucionais, como o Tópico #09 relacionado a “Compra de Vacinas por Presidente” que evidencia a personalização da responsabilidade sobre a vacinação, frequentemente culpabilizando ou defendendo figuras políticas, como o então presidente do Brasil, Jair Bolsonaro, ou o presidente dos Estados Unidos à época, Donald Trump. Já o Tópico #05 nomeado como “Fontes de notícias” agrupa mensagens que mencionam portais de mídia, muitas vezes desacreditando suas informações e acusando-os de manipulação, o que pode reforçar a desconfiança em fontes oficiais. Por fim, o Tópico #08 diz respeito à “Vacinação em Escolas”, onde a vacinação de professores e o retorno às aulas presenciais são discutidos em tom de preocupação, crítica ou resistência. Esses resultados mostram que, para além do

Tabela 2: Tabela com os principais tópicos identificados por meio do BERTopic.

#	Rótulo	#Msgs	Palavras
01	Mortes por vacina	1442	mortes, morreram, morreu, morte, morrer, morrendo, vacina, vacinados, vacinas, pessoas
02	Vacina Chinesa	1098	chinesa, vacina chinesa, china, vacina china, chinês, vacina, chineses, vírus, chinês, dória, doria
03	Coronavac	841	coronavac, contra coronavírus, coronavírus, vacina coronavac, vacina contra, abaixo redes, saiba link, novo coronavírus, contra novo, sociais vacinação
04	Vacinação Infantil	725	crianças, vacinação crianças, crianças anos, vacinação infantil, vacinar crianças, infantil, filhos, vacinar, vacina crianças, covid crianças
05	Fontes de notícias	566	ebc, jrmunews, mc, lendo jrmunews, lendo, cnn, anos, im, jp, fonte
06	Vacinação contra Gripe	540	gripe, contra gripe, influenza, vacinação contra, contra influenza, vacina gripe, campanha, vacinação, campanha, vacinação, contra
07	Doses de Vacina	532	dose, dose vacina, segunda dose, doses, segunda, doses vacina, duas doses, terceira dose, reforço, aplicação
08	Vacinação em Escolas	448	escolas, educação, professores, aulas, trabalhadores, educação, profissionais, educação, alunos, escola, vacinação, aulas presenciais
09	Compra de Vacinas por Presidente	433	presidente, vacina, comprar, presidente comprou, trump, vacinas, pra, dizer presidente, comprou, comprar vacina
10	Vídeos sobre Vacinas	429	vídeo, vídeos, assista, video, veja vídeo, assistir vídeo, assistir, veja, sobre vacinas, assista vídeo

Tabela 3: Exemplo de mensagem para cada um dos tópicos gerado pelo BerTopic.

Rótulo do Tópico	Exemplo de mensagem na base
Mortes por vacina	"fod* vai morrer muita gente, estão loucos para vacinar as pessoas"
Vacina Chinesa	"Os globalistas ganham muito dinheiro e reduzem a população com o vírus chinês e as vacinas."
Coronavac	"VAMOS COMEMORAR DIVULGUE "ESPECTACULAR NOTÍCIA" A vacina contra Corona vírus da Coronavac mostra eficácia de 95% e vacinação começa em Janeiro/21."
Vacinação Infantil	"Tem que prender a mãe que leva os filhos de menor de idade para tomar esse lixo de vacina . Essa bosta de vacina vai deixar sequelas nas crianças pelo resto da vida ."
Fontes de notícias	"MENTIRA NAO TEM VARIANTE A CNN TAO FAZENDO DO Povo PALHACO NAO ACRETIDEM MENTIRA INFORMACAO POR VARIOS MEDICOS ATE O PRESIDENTE DA AFRICANO DESMENTIU..."
Vacinação contra Gripe	"Confira pontos de vacinação contra Covid-19 e Influenza neste sábado..."
Doses de Vacina	"#IMUNIZAÇÃO #COVID-19 Crianças a partir de 3 anos podem tomar a primeira dose da vacina contra a Covid-19 a partir desta sexta-feira"
Vacinação em Escolas	"Saúde antecipa vacinação de trabalhadores de educação e autoriza imunização da população geral por idade"
Compra de Vacinas por Presidente	"Culpa do presidente q não comprou vacina quando foi oferecida, que não seguiu as orientações dos médicos do mundo todo"
Vídeos sobre Vacinas	"Olha assista este vídeo e vamos compaillhar o máximo a vacina para covid 19 dos Estados Unidos"

Tabela 4: Ranking dos 20 domínios mais compartilhados.

URL de Domínio	Ocorrências	Categoria	Associado à ANJ
gazetabrasil.com.br	10,414	Notícias	Não
youtu.be	4,471	Redes Sociais	—
chat.whatatsapp.com	3,170	Redes Sociais	—
t.me	3,056	Redes Sociais	—
www.boatos.org	2,400	Fact-checking	—
www.instagram.com	2,280	Redes Sociais	—
terrabrasilnoticias.com	2,199	Notícias	Não
www.facebook.com	2,096	Redes Sociais	—
twitter.com	1,753	Redes Sociais	—
g1.globo.com	1,133	Notícias	Sim
noticias.uol.com.br	930	Notícias	Sim
campos24horas.com.br	859	Notícias	Não
www.contrafatos.com.br	734	Notícias	Não
www1.folha.uol.com.br	675	Notícias	Sim
ootimista.com.br	636	Notícias	Sim
pubmed.ncbi.nlm.nih.gov	580	Artigos Científicos	—
serido360.com.br	531	Notícias	Não
www.cnnbrasil.com.br	474	Notícias	Não*
tribunacional.com.br	450	Notícias	Não
atrombeta.news.com.br	445	Notícias	Não

conteúdo estritamente sanitário, os discursos se entrelaçam com posicionamentos ideológicos, críticas à mídia e disputas políticas, compondo um ecossistema de desinformação multifacetado.

4.4 Análise das URLs

O WhatsApp integra-se a diversas fontes externas, e mensagens frequentemente recorrem a URLs como estratégia para conferir aparente credibilidade, mesmo em casos de desinformação. Para analisar essa prática, foi elaborada uma lista dos URLs mais compartilhados, a fim de identificar temas recurrentes e possíveis finalidades do compartilhamento. A Tabela 4 mostra que, entre as 20 URLs mais compartilhadas, 12 estão diretamente associadas a portais de notícias, demonstrando assim o comportamento esperado de utilizar links jornalísticos como forma de legitimar informações compartilhadas nos grupos. Alguns são amplamente reconhecidos pela credibilidade e rigor informativo atestados pela Associação Nacional de Jornais (ANJ)², ou reconhecidos internacionalmente (e.g., CNN). Por outro lado, há também veículos de menor confiabilidade, não vinculados à ANJ, conhecidos por apresentar viés editorial e, em certos casos, por disseminar desinformação ou distorcer dados para sustentar narrativas específicas [68, 69].

Além dos portais jornalísticos, essa análise também revela uma forte presença de domínios pertencentes a outras plataformas de redes sociais, como YouTube, Facebook, Twitter e Instagram, entre os links mais compartilhados. Dos 10 domínios mais frequentes, 6 pertencem a esse tipo de plataforma, evidenciando uma dinâmica de interconexão. Nesse ecossistema, o WhatsApp, devido ao seu grande alcance na população, funciona como uma espécie de vitrine ou repositório informal [70], em que conteúdos originalmente publicados em outras plataformas sociais de menor alcance são redistribuídos, muitas vezes até fora de seu contexto original, ampliando seu alcance e facilitando sua viralização.

Entre os 20 domínios mais compartilhados, destacam-se dois casos atípicos: Boatos.org e PubMed. A presença da PubMed³, uma das principais plataformas de artigos científicos da área biomédica,

indica o uso de referências técnicas para conferir credibilidade às mensagens. A citação de artigos científicos pode conferir elevada credibilidade às mensagens, o que aumenta o potencial persuasivo dessas referências. No entanto, uma análise qualitativa mais individual, indica que muitos estudos foram descontextualizados ou distorcidos para sustentar argumentos antivacina, conforme apontado em [71]. Já os links do Boatos.org, agência brasileira de verificação de fatos, sugerem tentativas de alguns usuários em desmentir rumores e promover contra-narrativas em meio a conteúdos predominantemente alarmistas e conspiratórios, esses esforços de *fact-checking* sugerem sinais de uma resistência discursiva ou tentativa de promover contra-narrativas em um ambiente frequentemente dominado por conteúdos alarmistas e conspiracionistas.

5 CONCLUSÃO

Este estudo analisou de forma abrangente a disseminação de (des)informação sobre vacinas em grupos públicos brasileiros no WhatsApp, no período de 2020 a 2023. Por meio de técnicas de Processamento de Linguagem Natural (PLN) aplicadas a milhares de mensagens, foi possível mapear as principais narrativas associadas à vacinação durante a pandemia de COVID-19, caracterizadas por conteúdos alarmistas, medo e desconfiança. Uma contribuição central da pesquisa foi a identificação e categorização temática do discurso antivacina, incluindo associações com mortes supostamente inexplicadas, desconfiança em relação a vacinas de origem chinesa e teorias conspiratórias sobre efeitos adversos. Essas narrativas mostraram-se fortemente vinculadas a posicionamentos ideológicos e políticos.

A análise evidenciou padrões de disseminação potencialmente coordenados, com mensagens replicadas por um número restrito de usuários altamente ativos, sugerindo estratégias deliberadas de amplificação da desinformação. Observou-se também uma distribuição assimétrica da atividade, em que poucos usuários concentraram a maior parte das postagens, enquanto a maioria contribuiu de forma pontual. Em síntese, este trabalho ofereceu uma caracterização detalhada de uma base de dados relevante durante um período crítico da nossa história recente. Direções para trabalhos futuros incluem explorar a detecção automática de desinformação relacionada à saúde em grupos do WhatsApp.

AGRADECIMENTOS

Este trabalho foi realizado com apoio financeiro do Centro de Inovação em Inteligência Artificial para a Saúde (CI-IA Saúde), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), e do Instituto Nacional de Ciência e Tecnologia em Inteligência Artificial Responsável para Linguística Computacional, Tratamento e Disseminação de Informação (INCT-TILD-IAR) – Grant #408490/2024-1.

²<https://www.anj.org.br/associados>

³<https://pubmed.ncbi.nlm.nih.gov/>

REFERÊNCIAS

- [1] UNICEF, "Immunization. access to life-saving vaccines to combat diseases is not universal and varies across regions and countries," UNICEF Data: Monitoring the situation of children and women, 2025, acessado em: 2025-07-14. [Online]. Available: <https://data.unicef.org/topic/child-health/immunization/>
- [2] M. Giaccherini, J. Kopinska, G. Rovigatti *et al.*, "Vax populi: the social costs of online vaccine skepticism," CESifo, Tech. Rep., 2022.
- [3] F. Benevenuto and P. Melo, "Misinformation campaigns through whatsapp and telegram in presidential elections in brazil," *Communications of the ACM*, vol. 67, no. 8, pp. 72–77, 2024.
- [4] P. Melo, C. C. Vieira, K. Garimella, P. O. V. de Melo, and F. Benevenuto, "Can whatsapp counter misinformation by limiting message forwarding?" in *International Conference on Complex Networks and Their Applications*, 2019, pp. 372–384.
- [5] J. C. S. Reis and F. Benevenuto, "Supervised learning for misinformation detection in whatsapp," in *Proceedings of the Brazilian Symposium on Multimedia and the Web*, 2021, p. 245–252.
- [6] C. Machado, B. Kira, V. Narayanan, B. Kollanyi, and P. Howard, "A study of misinformation in whatsapp groups with a focus on the brazilian presidential elections." in *Proc. of the World Wide Web Conference*, 2019, p. 1013–1019.
- [7] G. Resende, P. Melo, H. Sousa, J. Messias, M. Vasconcelos, J. Almeida, and F. Benevenuto, "(Mis)information Dissemination in WhatsApp: Gathering, Analyzing and Countermeasures," in *The World Wide Web Conference*, 2019, p. 818–828.
- [8] V. S. Bursztyn and L. Birnbaum, "Thousands of small, constant rallies: A large-scale analysis of partisan whatsapp groups," in *IEEE/ACM Int'l Conference on Advances in Social Networks Analysis and Mining*, 2019, pp. 484–488.
- [9] S. Vijaykumar, D. T. Rogerson, Y. Jin, and M. S. de Oliveira Costa, "Dynamics of social corrections to peers sharing covid-19 misinformation on whatsapp in brazil," *Journal of the American Medical Informatics Association*, vol. 29, no. 1, pp. 33–42, 11 2021.
- [10] D. A. Pinheiro, M. G. L. De Araújo, K. B. De Souza, B. Campos, E. De Oliveira, R. Lima, G. Ferreira, A. De Freitas, C. Toledo, G. De Souza *et al.*, "Sharing fake news about health in the cross-platform messaging app whatsapp during the covid-19 pandemic: A pilot study," *Int. J. Sci. Res. Manag.*, vol. 8, pp. 403–410, 2020.
- [11] N. A. Atehortua and S. Patino, "Covid-19, a tale of two pandemics: novel coronavirus and fake news messaging," *Health Promotion International*, vol. 36, no. 2, pp. 524–534, 01 2021.
- [12] A. Cavalini, T. Donadia, F. Malini, and G. Comarella, "Detecting misinformation on telegram anti-vaccine communities," in *Anais do XXXIX Simpósio Brasileiro de Bancos de Dados (SBBD)*, 2024, pp. 729–735.
- [13] G. S. Carlos, L. R. H. Maia, L. Massarani, T. Oliveira, and F. J. S. Maia, "Autoridade médica em grupos antivacina no whatsapp no brasil," *Cuadernos. info*, no. 60, pp. 143–167, 2025.
- [14] World Health Organization, "Ten threats to global health in 2019," <https://www.who.int/news-room/spotlight/ten-threats-to-global-health-in-2019>, 2019.
- [15] WHO, "Infodemic," World Health Organization, accessed em: 2025-07-14. [Online]. Available: https://www.who.int/health-topics/infodemic#tab=tab_1
- [16] I. J. Borges do Nascimento, A. B. Pizarro, J. M. Almeida, N. Azzopardi-Muscat, M. A. Gonçalves, M. Björklund, and D. Novillo-Ortiz, "Infodemics and health misinformation: a systematic review of reviews," *Bulletin of the World Health Organization*, vol. 100, no. 9, pp. 544–561, Sep, 2022, epub 2022 Jun 30.
- [17] I. Skafle, A. Nordahl-Hansen, D. S. Quintana, R. Wynn, and E. Gabarron, "Misinformation about covid-19 vaccines on social media: rapid review," *Journal of medical Internet research*, vol. 24, no. 8, p. e37367, 2022.
- [18] T. R. Nogueira, T. R. N. Dias, M. E. S. Rodrigues, É. L. Pereira, E. M. B. Ramos, and M. C. Delduque, "Vozes do desconhecimento: uma análise de discurso crítico dos movimentos antivacinas brasileiros no facebook," *Saúde em Debate*, vol. 49, p. e9849, 2025.
- [19] A. Nguyen and D. Catalan-Matamoros, "Anti-vaccine discourse on social media: An exploratory audit of negative tweets about vaccines and their posters," *Vaccines*, vol. 10, no. 12, p. 2067, 2022.
- [20] E. F. D'Almonte, E. L. Siqueira, G. de Araújo *et al.*, "Vacinias e desinformação: uma análise de conteúdo sobre fake news apuradas por plataformas de debunking em redes sociais," *Recis*, vol. 17, no. 3, 2023.
- [21] U. K. Ecker, S. Lewandowsky, J. Cook, P. Schmid, L. K. Fazio, N. Brashier, P. Kendeou, E. K. Vraga, and M. A. Amazeen, "The psychological drivers of misinformation belief and its resistance to correction," *Nature Reviews Psychology*, vol. 1, no. 1, pp. 13–29, 2022.
- [22] T. DeAngelis. (2023, jan) Psychologists are taking aim at misinformation with these powerful strategies. [Online]. Available: <https://www.apa.org/monitor/2023/01/trends-taking-aim-misinformation>
- [23] U. Voss, K. Schermelleh-Engel, L. Hauser, M. Holzmann, D. Fichtner, S. Seifert, A. Klimke, and S. Windmann, "Alike but not the same: Psychological profiles of covid-19 vaccine skeptics," *Health psychology open*, vol. 11, p. 20551029241248757, 2024.
- [24] L. Simione, M. Vagni, C. Gnagnarella, G. Bersani, and D. Pajardi, "Mistrust and beliefs in conspiracy theories differently mediate the effects of psychological factors on propensity for covid-19 vaccine," *Frontiers in psychology*, vol. 12, p. 683684, 2021.
- [25] A. Sapienza and R. Falcone, "The role of trust in covid-19 vaccine acceptance: considerations from a systematic review," *International journal of environmental research and public health*, vol. 20, no. 1, p. 665, 2022.
- [26] D. Vlasak, R. E. Dinero, and N. A. Roitman, "Vaccine hesitancy at both ends of the socioeconomic spectrum: a new paradigm for understanding the role of systemic inequity," *Journal of Public Health and Emergency*, vol. 7, no. 0, 2023.
- [27] M. Wróblewski and A. Melé, "Political polarization may affect attitudes towards vaccination: an analysis based on the european social survey data from 23 countries," *European Journal of Public Health*, vol. 34, no. 2, pp. 375–379, 01 2024.
- [28] E. Denniss and R. Lindberg, "Social media and the spread of misinformation: infectious and a threat to public health," *Health promotion international*, vol. 40, no. 2, p. daaf023, 2025.
- [29] J. C. Reis, P. Melo, M. Silva, and F. Benevenuto, "Desinformação em plataformas digitais: Conceitos, abordagens tecnológicas e desafios," *Sociedade Brasileira de Computação*, 2023.
- [30] D. Kbaier, A. Kane, M. McJury, and I. Kenny, "Prevalence of health misinformation on social media—challenges and mitigation before, during, and beyond the covid-19 pandemic: Scoping literature review," *J Med Internet Res*, vol. 26, p. e38786, Aug 2024.
- [31] F. Cascini, A. Pantovic, Y. A. Al-Ajlouni, G. Failla, V. Puleo, A. Melnyk, A. Lontano, and W. Ricciardi, "Social media and attitudes towards a covid-19 vaccination: A systematic review of the literature," *EClinicalMedicine*, vol. 48, 2022.
- [32] D. Kansaon, P. Melo, and F. Benevenuto, "'click here to join': A large-scale analysis of topics discussed by brazilian public groups on whatsapp," in *Proc.of the Brazilian Symposium on Multimedia and the Web*, ser. WebMedia'22, 2022.
- [33] J. C. Reis, P. Melo, K. Garimella, and F. Benevenuto, "Detecting misinformation on whatsapp without breaking encryption," *Association for the Advancement of Artificial Intelligence*, 2020.
- [34] P. F. Melo, M. Hoseini, S. Zannettou, and F. Benevenuto, "Don't break the chain: Measuring message forwarding on whatsapp," in *Proceedings of the International AAAI Conference on Web and Social Media*, 2024, pp. 1054–1067.
- [35] J. C. Reis, P. Melo, K. Garimella, and F. Benevenuto, "Can whatsapp benefit from debunked fact-checked stories to reduce misinformation?" *Harvard Kennedy School Misinformation Review*, 2020.
- [36] J. C. S. Reis, P. Melo, F. Belém, F. Murai, J. M. Almeida, and F. Benevenuto, "Helping fact-checkers identify fake news stories shared through images on whatsapp," in *Proc. of the Brazilian Symposium on Multimedia and the Web*, 2023, p. 159–167.
- [37] G. Resende, P. Melo, J. C. S. Reis, M. Vasconcelos, J. M. Almeida, and F. Benevenuto, "Analyzing textual (mis)information shared in whatsapp groups," in *Proc. of the 10th ACM Conference on Web Science*, 2019, p. 225–234.
- [38] L. Cesarino, "Como vencer uma eleição sem sair de casa: a ascensão do populismo digital no brasil," *Internet & sociedade*, vol. 1, no. 1, pp. 91–120, 2020.
- [39] J. C. Reis, P. Melo, K. Garimella, J. M. Almeida, D. Eckles, and F. Benevenuto, "A dataset of fact-checked images shared on whatsapp during the brazilian and indian elections," in *Proceedings of the international AAAI conference on web and social media*, 2020, pp. 903–908.
- [40] J. A. Caetano, G. Magno, M. Gonçalves, J. Almeida, H. T. Marques-Neto, and V. Almeida, "Characterizing attention cascades in whatsapp groups," in *Proceedings of the 10th ACM conference on web science*, 2019, pp. 27–36.
- [41] E. Alfaris, Y. Alhazzani, A. Alkhenizan, F. Ifran, N. Almoneef, N. Alyousefi, H. Alfaris, K. Alodhaibi, and A. M. Ahmed, "Assessing the validity of health messages used by the saudi public in whatsapp," *Patient Preference and Adherence*, vol. 17, pp. 67–73, 2023, pMID: 36632071.
- [42] V. J. Wirtz, G. Millán-Garduño, J. Hegewisch-Taylor, A. Dreser, A. Anaya-Sánchez, T. T. González-Vázquez, R. Escalera, and P. Torres-Pereda, "Misinformation messages shared via whatsapp in mexico during the covid-19 pandemic: an exploratory study," *Health Promotion International*, vol. 38, no. 3, p. daad041, 05 2023.
- [43] M. Winters, B. Oppenheim, P. Sengeh, M. B. Jaloh, N. Webber, S. A. Pratt, B. Leigh, H. Molsted-Alvesson, Z. Zeebari, C. J. Sundberg *et al.*, "Debunking highly prevalent health misinformation using audio dramas delivered by whatsapp: evidence from a randomised controlled trial in sierra leone," *BMJ global health*, vol. 6, no. 11, p. e006954, 2021.
- [44] C. P. Galhardi, N. P. Freire, M. C. M. Fagundes, M. C. d. S. Minayo, and I. C. K. O. Cunha, "Fake news and vaccine hesitancy in the covid-19 pandemic in brazil," *Ciência & Saúde Coletiva*, vol. 27, pp. 1849–1858, 2022.
- [45] S. S. Sundar, M. D. Molina, and E. Cho, "Seeing is believing: Is video modality more powerful in spreading fake news via online messaging apps?" *Journal of Computer-Mediated Communication*, vol. 26, no. 6, pp. 301–319, 08 2021.
- [46] P. Malhotra, "A relationship-centered and culturally informed approach to studying misinformation on covid-19," *Social Media + Society*, vol. 6, no. 3, p. 2056305120948224, 2020, pMID: 34192033.
- [47] D. Agarwal, F. Shahid, and A. Vashistha, "Conversational agents to facilitate deliberation on harmful content in whatsapp groups," *Proc. ACM Hum.-Comput. Interact.*, vol. 8, no. CSCW2, Nov. 2024.
- [48] A. E. Sharma, K. Khosla, K. Potharaju, A. Mukherjea, and U. Sarkar, "Covid-19-associated misinformation across the south asian diaspora: Qualitative study of whatsapp messages," *JMIR Infodemiology*, vol. 3, p. e38607, Jan 2023.

- [49] A. Puska, L. A. Baroni, M. C. Canal, L. S. G. Piccolo, and R. Pereira, "Whatsapp and false information: a value-oriented evaluation," in *Proc. of the Brazilian Symposium on Human Factors in Computing Systems*, 2020.
- [50] J. M. Maweu, "fake elections?" cyber propaganda, disinformation and the 2017 general elections in kenya," *African Journalism Studies*, vol. 40, no. 4, pp. 62–76, 2019.
- [51] K. Garimella and G. Tyson, "Whatapp doc? a first look at whatsapp public group data," in *Proceedings of the international AAAI conference on web and social media*, vol. 12, no. 1, 2018.
- [52] P. Melo, F. Benevenuto, D. Kansaon, V. Mafra, and K. Sá, "Monitor de whatsapp: Um sistema para checagem de fatos no combate à desinformação," in *Anais Estendidos do Simpósio Brasileiro de Sistemas Multimídia e Web*, 2021, pp. 79–82.
- [53] I. Claudio de Sá, J. M. Monteiro, J. M. Franco da Silva, L. M. Medeiros, P. J. C. Mourão, and L. C. Carneiro da Cunha, "Digital lighthouse: A platform for monitoring public groups in whatsapp," in *Proceedings of the 23rd International Conference on Enterprise Information Systems - Volume 1: ICEIS*, INSTICC. SciTePress, 2021, pp. 297–304.
- [54] P. Melo, J. Messias, G. Resende, K. Garimella, J. Almeida, and F. Benevenuto, "Whatsapp monitor: A fact-checking system for whatsapp," in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 13, 2019, pp. 676–677.
- [55] N. Reimers and I. Gurevych, "Sentence-bert: Sentence embeddings using siamese bert-networks," in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 11 2019. [Online]. Available: <https://arxiv.org/abs/1908.10084>
- [56] M. Douze, A. Guzhva, C. Deng, J. Johnson, G. Szilvassy, P.-E. Mazaré, M. Lomeli, L. Hosseini, and H. Jégou, "The faiss library," 2024.
- [57] M. Grootendorst, "Bertopic: Neural topic modeling with a class-based tf-idf procedure," *arXiv preprint arXiv:2203.05794*, 2022.
- [58] N. Reimers and I. Gurevych, "Making monolingual sentence embeddings multilingual using knowledge distillation," in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 11 2020. [Online]. Available: <https://arxiv.org/abs/2004.09813>
- [59] T. Sainburg, L. McInnes, and T. Q. Gentner, "Parametric umap embeddings for representation and semisupervised learning," *Neural Computation*, vol. 33, no. 11, pp. 2881–2907, 2021.
- [60] L. McInnes, J. Healy, and S. Astels, "hdbscan: Hierarchical density based clustering," *The Journal of Open Source Software*, vol. 2, no. 11, mar 2017.
- [61] G. Team, A. Kamath, J. Ferret, S. Pathak, N. Vieillard, R. Merhej, S. Perrin, T. Matejovicova, A. Ramé, M. Rivière, L. Rouillard, T. Mesnard, G. Cideron, J. bastien Grill, S. Ramos, E. Yvinec, M. Casbon, E. Pot, I. Penchev, G. Liu, F. Visin, K. Kenearly, L. Beyer, X. Zhai, A. Tsitsulin, R. Busa-Fekete, A. Feng, N. Sachdeva, B. Coleman, Y. Gao, B. Mustafa, I. Barr, E. Parisotto, D. Tian, M. Eyal, C. Cherry, J.-T. Peter, D. Sinopalnikov, S. Bhupatiraju, R. Agarwal, M. Kazemi, D. Malkin, R. Kumar, D. Vilas, I. Brusilovsky, J. Luo, A. Steiner, A. Friesen, A. Sharma, A. Sharma, A. M. Gilady, A. Goedekemeyer, A. Saade, A. Feng, A. Kolesnikov, A. Bendebury, A. Abdagic, A. Vadi, A. György, A. S. Pinto, A. Das, A. Bapna, A. Miech, A. Yang, A. Paterson, A. Shenoy, A. Chakrabarti, B. Piot, B. Wu, B. Shahriari, B. Petrini, C. Chen, C. L. Lan, C. A. Choquette-Choo, C. Carey, C. Brick, D. Deutsch, D. Eisenbud, D. Cattle, D. Cheng, D. Paparas, D. S. Sreepathihihalli, D. Reid, D. Tran, D. Zelle, E. Noland, E. Huizenga, E. Kharitonov, F. Liu, G. Amirkhanyan, G. Cameron, H. Hashemi, H. Klimczak-Plucińska, H. Singh, H. Mehta, H. T. Lehri, H. Hazimeh, I. Ballantyne, I. Szpektor, I. Nardini, J. Pouget-Abadie, J. Chan, J. Stanton, J. Wieting, J. Lai, J. Orbay, J. Fernandez, J. Newlan, J. yeong Ji, J. Singh, K. Black, K. Yu, K. Hui, K. Vodrahalli, K. Greff, L. Qiu, M. Valentine, M. Coelho, M. Ritter, M. Hoffman, M. Watson, M. Chaturvedi, M. Moynihan, M. Ma, N. Babar, N. Noy, N. Byrd, N. Roy, N. Momchev, N. Chauhan, N. Sachdeva, O. Bunyan, P. Botarda, P. Caron, P. K. Rubenstein, P. Culliton, P. Schmid, P. G. Sessa, P. Xu, P. Stanczyk, P. Tafti, R. Shivanma, R. Wu, R. Pan, R. Rokni, R. Willoughby, R. Vallu, R. Mullins, S. Jerome, S. Smoot, S. Girgin, S. Iqbal, S. Reddy, S. Sheth, S. Pöder, S. Bhatnagar, S. R. Panyam, S. Eiger, S. Zhang, T. Liu, T. Yacovone, T. Liechty, U. Kalra, U. Evcı, V. Misra, V. Roseberry, V. Feinberg, V. Kolesnikov, W. Han, W. Kwon, X. Chen, Y. Chow, Y. Zhu, Z. Wei, Z. Egyed, V. Cotruta, M. Giang, P. Kirk, A. Rao, K. Black, N. Babar, J. Lo, E. Moreira, L. G. Martins, O. Sanseviero, L. Gonzalez, Z. Gleicher, T. Warkentin, V. Mirrokni, E. Senter, E. Collins, J. Barral, Z. Ghahramani, R. Hadsell, Y. Matias, D. Sculley, S. Petrov, N. Fiedel, N. Shazeer, O. Vinyals, J. Dean, D. Hassabis, K. Kavukcuoglu, C. Farabet, E. Buchatskaya, J.-B. Alayrac, R. Anil, Dmitry, Lepikhin, S. Borgeaud, O. Bachem, A. Joulin, A. Andreev, C. Hardin, R. Dadashi, and L. Huszenot, "Gemma 3 technical report," 2025. [Online]. Available: <https://arxiv.org/abs/2503.1978>
- [62] Y. Tang, P. Ma, B. Kong, W. Ji, X. Gao, and X. Peng, "Espa: A novel approach for cross-platform event dissemination trend analysis between social network and search engine," in *Web Information Systems Engineering – WISE 2016*, W. Cellary, M. F. Mokbel, J. Wang, H. Wang, R. Zhou, and Y. Zhang, Eds. Cham: Springer International Publishing, 2016, pp. 489–504.
- [63] Ministério da Saúde, "Anvisa concede primeiro registro definitivo para vacina contra a Covid-19 nas Américas," Saúde e Vigilância Sanitária – gov.br, 2021, acessado em 21/07/2021. [Online]. Available: <https://www.gov.br/pt-br/noticias/saude-e-vigilancia-sanitaria/2021/02/anvisa-concede-primeiro-registro-definitivo-para-vacina-contra-a-covid-19-nas-americas>
- [64] R. Rodrigues, "Não tenham medo, diz Mônica Calazans, 1ª pessoa a ser vacinada no Brasil," Globo G1 SP, 2021, acessado em 21/07/2021. [Online]. Available: <https://g1.globo.com/sp/sao-paulo/noticia/2021/01/17/nao-tenham-medo-diz-monica-calazans-1a-pessoa-a-ser-vacinada-no-brasil.ghtml>
- [65] F. Brito, "Retrospectiva 2021: as milhões de vacinas Covid-19 que trouxeram esperança para o Brasil," Ministério da Saúde – gov.br, 2021, acessado em 21/07/2021. [Online]. Available: <https://www.gov.br/saude/pt-br/assuntos/noticias/2021/dezembro/retrospectiva-2021-as-milhoes-de-vacinas-covid-19-que-trouxeram-esperanca-para-o-brasil>
- [66] Ministério da Saúde, "Anvisa aprova vacina da Pfizer contra Covid para crianças de 5 a 11 anos," Saúde e Vigilância Sanitária – gov.br, 2021, acessado em 21/07/2021. [Online]. Available: <https://www.gov.br/anvisa/pt-br/assuntos/noticias-anvisa/2021/anvisa-aprova-vacina-da-pfizer-contra-covid-para-criancas-de-5-a-11-anos>
- [67] H. Rufino de Sousa Neto and L. Frezza Pisa, "Vacinação infantil da covid-19 e desinformação: A contribuição dos discursos políticos negacionistas para a proliferação da infodemia," *Revista Eixos Tech*, vol. 11, no. 4, jun. 2024.
- [68] J. M. Couto, J. C. Reis, I. Cunha, L. Araújo, and F. Benevenuto, "Caracterizando websites de baixa credibilidade no brasil," in *Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos (SBRC)*, 2022, pp. 503–516.
- [69] J. M. Couto, J. C. Reis, and F. Benevenuto, "Can computer network attributes be useful for identifying low-credibility websites? a case study in brazil," *Social Network Analysis and Mining*, vol. 14, no. 1, p. 153, 2024.
- [70] E. V. Pereira, P. Melo, M. Júnior, V. O. Mafra, J. C. S. Reis, and F. Benevenuto, "Analyzing youtube videos shared on whatsapp and telegram political public groups," in *Proceedings of the Brazilian Symposium on Multimedia and the Web*, 2022, p. 28–37.
- [71] M. Hameleers, "Disinformation as a context-bound phenomenon: toward a conceptual clarification integrating actors, intentions and techniques of creation and dissemination," *Communication Theory*, vol. 33, no. 1, pp. 1–10, 10 2022.