

## Vídeos sobre vacinas: quais fatores influenciam em maior visualização no YouTube?

Videos about vaccines: what factors influence on having higher views on YouTube?

Videos sobre vacunas: ¿qué factores influyen en una mayor visualización en YouTube?

*Arthur da Silva Lopes*<sup>1,a</sup>

[arthur.lopes@ufba.br](mailto:arthur.lopes@ufba.br) | <https://orcid.org/0000-0001-9137-3184>

*Antonio Marcos Pereira Brotas*<sup>2,b</sup>

[brotas@bahia.fiocruz.br](mailto:brotas@bahia.fiocruz.br) | <http://orcid.org/0000-0001-8438-2445>

<sup>1</sup> Universidade Federal da Bahia, Instituto de Saúde Coletiva. Salvador, BA, Brasil.

<sup>2</sup> Fundação Oswaldo Cruz, Centro de Pesquisa Gonçalo Moniz. Salvador, BA, Brasil.

<sup>a</sup> Graduação em Saúde pela Universidade Federal da Bahia.

<sup>b</sup> Doutorado em Cultura e Sociedade pela Universidade Federal da Bahia.

### RESUMO

Considerando-se a crescente importância do YouTube como fonte para busca de informações em saúde, o objetivo deste trabalho é analisar os fatores associados a um maior número de visualizações de vídeos sobre vacinas contra a covid-19. Para isso, usaram-se técnicas de Processamento de Linguagem Natural e modelagem estatística com base em 13.619 vídeos, abrangendo três tipos de variáveis: métricas gerais, conteúdo textual dos títulos e informações sobre os participantes dos vídeos. Entre os resultados, destacam-se os vídeos de duração média ou longa, postados durante a madrugada e nos fins de semana, com *tags*, descrição e títulos curtos, além de elementos controversos e presença de figuras masculinas e brancas em miniaturas. Os achados contribuem para uma melhor compreensão dos possíveis fatores a serem considerados na produção de conteúdo de comunicação em saúde sobre vacinas no YouTube.

**Palavras-chave:** Comunicação em saúde; Educação em saúde; YouTube; Vacinas; Engajamento.

## ABSTRACT

Considering the growing importance of YouTube as a source for health information search, the aim of this study was to analyze the factors associated with a higher number of views in videos about covid-19 vaccines. For this purpose, Natural Language Processing techniques and statistical modeling were employed based on 13,619 videos, encompassing three types of variables: general metrics, textual content of titles, and information about the participants in the videos. Among the results, videos of medium or long duration, posted during late hours and on weekends, with tags, descriptions, and short titles, along with controversial elements and the presence of male and white figures in thumbnails stand out. These findings contribute to a better understanding of the potential factors to be considered in the production of health communication content about vaccines on YouTube.

**Keywords:** Health communication; Health education; YouTube; Vaccines; Engagement.

## RESUMEN

Teniendo en cuenta la creciente importancia de YouTube como fuente de búsqueda de información en salud, el objetivo de este artículo es analizar los factores asociados a un mayor número de visualizaciones en videos sobre vacunas contra el covid-19. Para eso, se emplearon técnicas de Procesamiento del Lenguaje Natural y modelado estadístico basadas en 13,619 videos, que abarcan tres tipos de variables: métricas generales, contenido textual de títulos y información sobre los participantes en los videos. Entre los resultados, destacan los videos de duración media o larga, publicados durante altas horas de la noche y los fines de semana, con *tags*, descripciones y títulos cortos, junto con elementos controvertidos y la presencia de figuras masculinas y blancas en las miniaturas. Estos hallazgos contribuyen a una mejor comprensión de los posibles factores a tener en cuenta en la producción de contenido de comunicación de salud sobre vacunas en YouTube.

**Palabras clave:** Comunicación en salud; Educación en salud; YouTube; Vacunas; Interacción.

---

## INFORMAÇÕES DO ARTIGO

### Contribuição dos autores:

Concepção e desenho do estudo: Arthur da Silva Lopes.  
Aquisição, análise ou interpretação dos dados: Arthur da Silva Lopes.  
Redação do manuscrito: Arthur da Silva Lopes.  
Revisão crítica do conteúdo intelectual: Antonio Marcos Pereira Brotas.

**Declaração de conflito de interesses:** não há.

**Fontes de financiamento:** não houve.

**Considerações éticas:** não há.

**Agradecimentos/Contribuições adicionais:** não há.

**Histórico do artigo:** submetido: 28 jun. 2023 | aceito: 29 ago. 2023 | publicado: 28 mar. 2024.

**Apresentação anterior:** não houve.

**Licença CC BY-NC atribuição não comercial.** Com essa licença é permitido acessar, baixar (*download*), copiar, imprimir, compartilhar, reutilizar e distribuir os artigos, desde que para uso não comercial e com a citação da fonte, conferindo os devidos créditos de autoria e menção à Reciis. Nesses casos, nenhuma permissão é necessária por parte dos autores ou dos editores.

## INTRODUÇÃO

André Pereira Neto e colaboradores (2015) apontam que a internet pode ser considerada a mais emblemática Tecnologia da Informação e Comunicação (TIC) da contemporaneidade. Ela provocou mudanças substanciais nas maneiras como os indivíduos se informam, se comunicam e se relacionam. A internet reconfigurou os papéis daqueles que produzem e daqueles que consomem informações, rompeu com as fronteiras geográficas e tornou ubíquo e instantâneo seu acesso e compartilhamento.

Acerca do impacto dessas transformações no campo da saúde, há de se destacar aquelas inseridas no âmbito da comunicação e no âmbito da educação, sobretudo em razão do que McClung, Murray e Heitlinger apontaram no final do século XX: “[...] o público quer ter acesso à informação sobre problemas médicos. Os fornecedores de informação na internet (net) responderam a este desejo, e a informação médica é abundante” (1998, p. 01, tradução nossa).

Naquela época, essas informações estavam circunscritas aos motores de busca e *websites*. Atualmente, as pessoas podem encontrar estabelecimentos e profissionais de saúde, acessar as avaliações de outras pessoas sobre o atendimento em *sites* médicos, informar-se acerca de possibilidades terapêuticas e diagnósticas, assim como discutir entre pares em fóruns e redes sociodigitais (Garbin; Pereira Neto; Guilam, 2008).

## INFORMAÇÕES SOBRE SAÚDE EM PLATAFORMAS DE MÍDIA SOCIAL: O CASO DO YOUTUBE

O Digital News Report (Newman, 2022) – relatório publicado pela agência Reuters –, aponta que as fontes de informação se encontram notadamente em ambientes *on-line* (83%), sendo o YouTube aquela mais utilizada para consumo de notícias (43%), e a segunda para usos gerais (77%), atrás apenas do WhatsApp (78%). É por essa razão que o YouTube é tratado aqui como lócus de investigação e os vídeos produzidos nessa plataforma consistem em nosso objeto de análise.

Não obstante a disponibilidade e a abundância de informações, aquilo que será consumido pelas pessoas nem sempre é confiável. Fialho *et al.* (2021), ao analisarem 242 vídeos em língua portuguesa sobre infarto agudo do miocárdio, identificaram que um terço dos vídeos eram irrelevantes, continham dados imprecisos, e entre os relevantes a qualidade DISCERN<sup>1</sup> média foi baixa ou moderada.

Em um estudo mais recente, realizado por Dalpoz *et al.* (2022), no qual foram investigados 48 vídeos educacionais sobre prevenção da cárie dental, constatou-se que a maioria dos vídeos tinha qualidade moderada de acordo com o critério DISCERN. No que tange à confiabilidade (JAMA), no geral, ela foi baixa por falta de transparência quanto à proveniência das informações apresentadas.

Chan e colaboradores (2021), por sua vez, analisaram vídeos sobre vacinas contra covid-19. A maior parte dos vídeos foi considerada não educacional, com qualidade DISCERN média “justa”. Já os vídeos educacionais tiveram a qualidade variando de “justa” a “excelente”, dependendo da origem do produtor. A confiabilidade da informação (HONcode<sup>2</sup>) foi “baixa” para a maioria dos vídeos não educacionais; para os educacionais, foi “moderada”. No geral, os autores do estudo sublinham que a qualidade e a confiabilidade das informações estão muito aquém do ideal.

A dialética entre qualidade e quantidade certamente se acirrou no contexto de infodemia e de desordem informacional (Wardle; Derakhshan, 2017). No que concerne às vacinas, o cenário se mostra propício para a atuação de movimentos antivacina, importantes agentes na promoção da hesitação vacinal (Pierri *et al.*, 2022), que sabemos ser hoje um problema relevante de saúde pública (WHO, 2019).

<sup>1</sup> Instrumento utilizado para avaliar a qualidade de informações relacionadas à saúde.

<sup>2</sup> Trata-se de uma certificação usada para avaliar a qualidade de informações relacionadas à saúde, por meio da verificação da procedência, da transparência e da intencionalidade (dimensão ética) do conteúdo produzido ou disponibilizado na internet.

Frente a esse cenário, pesquisadores têm se debruçado sobre o crescimento da desinformação acerca das vacinas no YouTube (Brotas; Costa; Massarani, 2021; Calvo; Cano-Orón; Llorca-Abad, 2022; Massarani *et al.*, 2021). Visando identificar a ocorrência de conteúdo desinformativo e compreender em quais situações ela se dá, esses trabalhos foram, em sua maioria, qualitativos e tomaram os vídeos selecionados como os mais relevantes enquanto objetos de análise.

É importante sublinhar que essa “relevância” é dada segundo os critérios do próprio YouTube, que, pertencendo a uma empresa privada, a Google, busca o maior número de interações (Souza, 2020). Trata-se de um sistema de recomendações que funciona com base nos *footprints* deixados pelos(as) usuários(as) ao utilizarem a plataforma. Apesar de serem esses os vídeos com maior probabilidade de acesso e de se tornarem fontes de informação, eles não representam a totalidade das produções disponíveis no YouTube sobre o tema.

Malgrado a importância e a relevância dos estudos qualitativos, estes ainda carecem estudos de caráter quantitativo que visem investigar quais os fatores associados a um maior número de visualizações de vídeos sobre vacinas no YouTube, considerando-se um conjunto compreensivo de variáveis possivelmente explicativas.

É importante destacar que pode haver outras produções de melhor qualidade, mas que ‘por alguma razão’ não conseguiram angariar audiência suficiente para disputar relevância com os demais. A literatura sobre o tema ainda não compreende o que leva certos vídeos sobre vacinas a terem mais audiência que outros. Nesse sentido, é fundamental entender os princípios que orientam a produção desses conteúdos – se são confiáveis ou não – e evidenciar possíveis critérios usados pelos sistemas de recomendação que favoreçam a desordem informacional.

Como Li e colaboradores (2020) destacam, o YouTube tem o potencial de ser uma ferramenta oportuna para a comunicação em saúde em situações de crise sanitária, tal qual a pandemia de covid-19, caso seja devidamente utilizada por instituições públicas e por profissionais de saúde para disseminar informações confiáveis e de qualidade. Nesse sentido, os autores recomendam que “[...] esses grupos devem encontrar estratégias para aumentar o número de visualizações e o impacto de seus vídeos [...]” na plataforma (Li *et al.*, 2020, p. 5, tradução nossa).

Destarte, tem-se como objetivo geral identificar os fatores associados a um maior número de visualizações para vídeos sobre vacina no YouTube, a fim de preencher a lacuna na literatura sobre a questão. Essas informações podem orientar profissionais e instituições de saúde e de divulgação científica, para que consigam competir por maiores audiências, fornecendo, assim, informações confiáveis e combatendo a desinformação.

## METODOLOGIA

De forma a responder ao objetivo da pesquisa, utilizou-se a abordagem de investigação quantitativa. A unidade de análise não será os vídeos com maior engajamento (como nos trabalhos qualitativos discutidos anteriormente), mas todos aqueles postados na plataforma entre 2020 e 2021, contendo a palavra “vacina” no título ou em sua descrição, voluntariamente publicados pelos usuários e passíveis de serem coletados através da plataforma YouTube Data Tools (Rieder, 2023).

Como resultado do procedimento de coleta, tem-se um total de 13.619 vídeos. Apesar de não tratarem do tema saúde, Munaro *et al.* (2021) e Yang *et al.* (2022) realizaram estudos estatísticos acerca dos fatores que influenciam em maiores visualizações e demais métricas no YouTube. As variáveis independentes são abrangentes, e foram selecionadas aquelas apontadas por esses trabalhos como tendo efeito estatisticamente significativo, quando aplicável, no número de visualizações (variável dependente deste estudo). Assim, estão discriminadas na Tabela 1.

**Tabela 1 – Descrição das variáveis categóricas relacionadas às métricas gerais**

Variável categórica	Notação	Categoria	N.	Proporção (%)
Postagem (parte do dia)	part_of_day	noite	5.280	38,76
		tarde	4.595	33,73
		madrugada	2.009	14,75
		manhã	1.735	12,73
Postagem (dia da semana)	day_of_week	segunda (seg)	2.640	19,38
		terça (ter)	2.586	18,98
		quarta (qua)	2.542	18,66
		quinta (qui)	2.396	17,59
		sexta (sex)	2.129	15,63
		sábado (sáb)	836	6,13
		domingo (dom)	490	3,59
		Duração do vídeo	duration	curta (< 10 min)
média (10-19:59 min)	1.085			79,66
longa (> 20 min)	824			6,05
Número de palavras no título	len_video_title	acima (>=12 palavras)	6.588	48,37
		abaixo (<= 10 palavras)	5.380	39,50
		média (10-12 palavras)	1.651	12,12
Número de caracteres caixa-alta	n_upper	abaixo (<= 6 caracteres)	10.428	76,65
		acima (>= 8 caracteres)	2.739	20,11
		média (6-8 caracteres)	452	3,31
Uso de descrição	desc_use	1 (Sim)	12.874	94,52
		0 (Não)	745	5,47

Fonte: Elaborada pelos autores.

## VARIÁVEIS ASSOCIADAS AO CONTEÚDO TEXTUAL DOS TÍTULOS

Para o Reconhecimento de Entidade Nomeada (NER, do inglês Named Entity Recognition) nos títulos, fez-se uso do algoritmo de autoria própria, o ProfileNER-classifier.<sup>3</sup> O NER é uma técnica pertencente à área de Processamento de Linguagem Natural (PLN) que compreende a identificação e a classificação de entidades, que podem ser qualquer conjunto de palavras que faça referência a uma determinada categoria.

Assim, dor de cabeça, tosse, dor muscular e febre são termos que compõem a entidade ‘síntomas’, por exemplo. Entre as entidades usadas neste estudo, temos: “vacinas”, “vacinas política” (vac\_política), “produtos”, “medicamentos”, “doenças”, “ciência”, “partes do corpo” (partes\_corpo), “eventos frequentes” (eventos\_freq) e “eventos raros” (eventos\_raros). A variável relacionada ao NER corresponde ao número de ocorrências de cada uma das entidades.

A entidade “vacinas” refere-se aos nomes dos laboratórios que as produzem (e.g. AstraZeneca, Oxford, Butantan, Fiocruz), enquanto “vacinas política” refere-se às designações: vachina, vacina do Doria, vacina da China, vacina Xing etc. Já a entidade “produtos” refere-se, dentre outros à: respiradores, seringa, máscara, EPI, PCR e leito.

Para a tarefa de detecção automatizada de conteúdo potencialmente desinformativo ou controverso nos títulos, fez-se uso de um modelo de aprendizado de máquina treinado inicialmente, conforme explicitam Lopes e Brotas (2023). Ao longo desta investigação, houve investimento na melhoria da *performance* do

<sup>3</sup> Algoritmo cujo funcionamento se dá através de dicionários de termos relacionados a cada entidade, quando na função NER, ou relacionados a cada perfil. Esses dicionários estão disponíveis na pasta [Profileclassifier](#).

modelo, nomeadamente a nível da seleção e da engenharia de atributos, e incremento nos dados de treino (conteúdo textual ilustrado na Figura 1), por meio da coleta e da classificação de *tweets*.



Figura 1 – Nuvem de palavras presentes nos dados de treino do modelo  
 Fonte: Elaborada pelos autores.

A escolha pelo Twitter se deu em função de dois aspectos: 1) adequação à tarefa a ser realizada pelo modelo; e 2) eficiência. Acerca do primeiro, o Twitter foi efetivamente palco de debates e disputas em torno das vacinas contra a covid-19 (Penteado *et al.*, 2021). No que concerne ao segundo, a plataforma é um espaço por meio do qual as pessoas e as instituições compartilham informações, opiniões, notícias, relatos cotidianos etc. em textos relativamente curtos (280 caracteres), reduzindo o custo operacional (da leitura manual para validação) e o computacional (no treinamento do modelo).

Os descritores usados na coleta foram correspondentes às categorias temáticas atribuídas, e a validação consistiu na leitura manual de todos os *tweets*, à exceção daqueles pertencentes aos temas “segurança” e “desinformação nomeada”, pois foi realizada uma amostragem de 50% dos *tweets* a eles designados. O resultado está sintetizado no Quadro 1.

**Quadro 1 – Temas dos novos dados de treino incorporados ao modelo**

Tema (n.)	Tweet-exemplo
Segurança (1.200)	A Anvisa DEVE, antes de recomendar, apresentar estudos robustos comprovando a eficácia e segurança da vacina. Oferecer uma 'vacina' sem comprovação de segurança é GENOCÍDIO.
Vacinas vencidas (300)	E faz alguma diferença aplicar “vacina” vencida? Arma biológica tem data de vencimento?
Vacina enquanto objeto de lucro (300)	Após a Pfizer lucrar trilhões com uma vacina bem questionável, e após o experimento de controle social concluído, declaramos fim da pandemia.
Efeitos adversos da vacina (300)	Essa mulher enlouqueceu de vez? OMS declarando fim da pandemia, países suspendendo a vacinação por conta dos efeitos adversos, estudos vindos à tona e concluindo que a vacina causa miocardite e ela vem falar em “intensificar”!?
Vaccine Adverse Event Reporting System (VAERS) (600)	Ela morreu dessa vacina, sim, TTP é umas das reações dessas vacinas, eles estão mentindo na cara dura, um bebê de 5 meses morreu lá nos Estados Unidos após a mãe tomar a vacina, está no VAERS.
Desinformação nomeada (conteúdo falso, enganoso e mentira) (1.843)	Publicações falsas sobre vacinas contra covid-19 cresceram 383% em 2 meses. Grupos antivacina usam redes sociais para dar eco a conteúdo falso e gerar desconfiança sobre futura campanha de vacinação.

Fonte: Elaborado pelos autores.

Dessa forma, houve melhora em todas as métricas de *performance*: acurácia dos dados de teste ainda não vistos pelo modelo (92% para 94%); média macro do *F1-score* (média harmônica entre precisão e

sensibilidade, 92% para 94%); precisão (92% para 95%); e sensibilidade (92% para 93%), apresentando melhor capacidade de generalização.

## VARIÁVEIS ASSOCIADAS AO MEDIADOR DOS VÍDEOS (AQUELE QUE APARECE NOS VÍDEOS)

Foi detectada a presença de rostos humanos nas *thumbnails* (miniaturas) dos vídeos, sob a hipótese de que a presença dos próprios mediadores na ‘capa’ de apresentação dos vídeos seja uma possível estratégia de mobilização do seu capital social (Recuero; Soares, 2021).

Foi realizada a classificação de gênero e raça com base nas *thumbnails* dos vídeos para que fosse possível investigar se ordenamentos decorrentes dessas categorias analítico-estruturais também tinham efeito no número de visualizações, usando o modelo de aprendizado de máquina e profundo, o VGG-Face (acurácia de 89,3%,  $F1 = 88,3\%$ ) através do *framework* DeepFace (Serengil; Ozpinar, 2020)

Tal qual discutido por Butler (2003), o próprio sexo seria um constructo sociocultural, uma vez que se faz possível percorrer sua genealogia e encontrar sentidos e significados atribuídos e diferentes daquele tido como hegemônico (biomédico). No entanto, a variável manteve-se nomeada como gênero, pois o sexo, em sua concepção dismórfico-anatômica (base na qual o modelo usado foi treinado), é compreendido como um componente constituinte da inteligibilidade (cis)gênero.

Para o gênero, o modelo possui 97% de acurácia. A maior parte foi designada à categoria “none”, ou seja, aquela em que o algoritmo não conseguiu detectar presença de rosto humano; logo, essas ocorrências foram desconsideradas, assim como aquelas em que não se pôde distinguir entre “homem” ou “mulher”.

No que concerne à raça/cor, o pressuposto teórico que amparou a adoção através da classificação automatizada foi o de Oracy Nogueira (2007), que aponta que o preconceito de raça no Brasil é mediado (e produzido) pela marca, ou seja, pelo fenótipo. No entanto, a complexidade do racismo brasileiro, e a própria noção de raça, evidenciada pelas disputas em torno das políticas públicas de ações afirmativas (Nunes, 2018) apresenta-se enquanto um desafio à tentativa de classificação automatizada, tal qual proposta inicialmente pelos autores do *DeepFace* e pelo banco de dados usado para o treinamento (Kärkkäinen; Joo, 2019).

Por essa razão, as categorias do modelo – baseadas no censo estadunidense: branco, preto, latino-americano, indiano, árabe e asiático – foram adaptadas, e a classificação inicialmente multiclasse tornou-se binária, tendo como variáveis: branco e não branco. Isso se deu em razão da inviabilidade teórica de simplesmente importar uma interpretação racial decorrente de mecanismo de manifestação social divergente (de origem, segundo Nogueira, no caso dos Estados Unidos), e de o objetivo desta investigação se restringir à identificação de possíveis assimetrias associadas aos diferentes processos de socialização mediados pela raça (racialização) de indivíduos fenotipicamente brancos, e daqueles que divergem desta norma.

O procedimento de classificação para raça/cor foi também alterado. Como o modelo possui 68% de acurácia, diferentes versões de imagem<sup>4</sup> foram utilizados para capturar múltiplos *frames* de *thumbnail* (pois em sua maioria são dinâmicos, quer dizer, apresentam fragmentos dos vídeos).

Assim, criou-se um *script* capaz de detectar a raça/cor aplicada a cada um dos *frames* (superando, assim, possíveis vieses associados à iluminação ou à posição da pessoa), e de modo que os resultados fossem armazenados em uma variável do tipo lista. Concluída essa etapa, a classificação final foi decorrente da moda (usando o módulo Statistics do Python, que indica o valor mais frequente de um determinado conjunto de dados) dessa lista. As estatísticas descritivas para presença de rosto humano, gênero e raça/cor estão dispostas na Tabela 2.

<sup>4</sup> O acesso às miniaturas dos vídeos no YouTube é feito com base na seguinte estrutura: <https://img.youtube.com/vi/<inserir-id-do-video-aqui>/inserir-formatos.jpg>; sendo os formatos utilizados aqui: o.jpg, hq1.jpg, hq2.jpg e hq3.jpg.

**Tabela 2 – Descrição das variáveis categóricas associadas ao mediador dos vídeos**

Variável categórica	Notação	Categoria	N.	Proporção (%)
Rosto na thumbnail	Rosto	0 (Não)	7.958	58,43
		1 (Sim)	5.661	41,56
Raça/etnia*	Raça/etnia	0 (Não branco)	163	3,01
		1 (Branco)	5.250	96,99
Gênero*	Gênero	0 (Mulher)	1.503	27,77
		1 (Homem)	3.910	72,23

Fonte: Elaborada pelos autores.

\* Para rosto “sim”. O total equivale a 5.413, pois subtraíram-se de 5.661 os vídeos em que o *script* para raça não pôde chegar a um resultado.

A análise da correlação entre as variáveis foi feita por meio da regressão binomial negativa, assim como nos trabalhos de Munaro *et al.* (2021) e Yang *et al.* (2022), em função de a natureza da variável dependente ser contável e apresentar sobre dispersão.

## RESULTADOS

Destarte, os resultados concernentes às variáveis nativas da plataforma estão dispostos na Tabela 3. A maior parte dos vídeos possuíam descrição, variável (*desc\_use*) que apresentou associação positiva com um maior número de visualizações na plataforma. Já acerca da duração do vídeo, a maioria daqueles sobre vacina foram curtos (< 10 min.), mas foram os de média ( $\beta = 1.22$ ) e longa ( $\beta = 1.02$ ) duração aqueles que mais estiveram associados a maior número de *views*.

**Tabela 3 – Descrição dos resultados do modelo concernente às variáveis associadas à plataforma**

Categoria	Coeficiente	p-valor ( $\alpha < 0,05$ )	(continua)	
			Intervalo de confiança (95%)	
intercept	2,5717	0,000	2,520	2,623
desc_use	2,1008	0,000	2,025	2,177
duration_curta	0,3336	0,000	0,299	0,368
duration_longa	1,0191	0,000	0,964	1,074
duration_media	1,2190	0,000	1,169	1,269
n_upper2_abaixo	0,9697	0,000	0,932	1,007
n_upper2_acima	1,0209	0,000	0,978	1,063
n_upper2_media	0,5811	0,000	0,512	0,650
len_video_title2_abaixo	0,9239	0,000	0,894	0,954
len_video_title2_acima	0,7911	0,000	0,761	0,821
len_video_title2_media	0,8567	0,000	0,817	0,896
len_tags2_abaixo	0,9026	0,000	0,872	0,934
len_tags2_acima	1,6691	0,000	1,636	1,702
part_of_day_madrugada	1,1820	0,000	1,141	1,223
part_of_day_manhã	0,3798	0,000	0,339	0,421
part_of_day_noite	0,5117	0,000	0,482	0,541
part_of_day_tarde	0,4982	0,000	0,468	0,528
videoCategoryLabel_Comedy	1,3932	0,000	0,996	1,791



(conclusão)

<b>Categoria</b>	<b>Coefficiente</b>	<b>p-valor (<math>\alpha &lt; 0,05</math>)</b>	<b>Intervalo de confiança (95%)</b>	
videoCategoryLabel_Education	1,3675	0,000	1,221	1,514
videoCategoryLabel_Entertainment	-0,0446	0,506	-0,176	0,087
videoCategoryLabel_Film	-2,0613	0,000	-2,409	-1,713
videoCategoryLabel_Gaming	-2,2304	0,000	-3,282	-1,178
videoCategoryLabel_Howto	1,3027	0,000	0,950	1,656
videoCategoryLabel_Music	4,6596	0,000	4,347	4,972
videoCategoryLabel_News	0,2837	0,000	0,157	0,410
videoCategoryLabel_Nonprofit	0,5720	0,000	0,293	0,851
videoCategoryLabel_People	-0,2379	0,001	-0,373	-0,103
videoCategoryLabel_Science	1,5568	0,000	1,407	1,707
videoCategoryLabel_Sports	0,1173	0,453	-0,189	0,424
videoCategoryLabel_Travel	-4,1069	0,000	-4,928	-3,286
day_of_week_dom	0,4412	0,000	0,362	0,520
day_of_week_qua	0,1393	0,000	0,100	0,179
day_of_week_quin	0,1582	0,000	0,120	0,197
day_of_week_sab	0,5492	0,000	0,487	0,612
day_of_week_seg	0,4288	0,000	0,386	0,471
day_of_week_sex	0,5040	0,000	0,464	0,544
day_of_week_ter	0,3510	0,000	0,312	0,390

Fonte: Elaborada pelos autores.

No que concerne ao número de palavras nos títulos, tem-se maior ocorrência de vídeos acima da média ( $\geq 12$  palavras), mas aqueles com 10 palavras ou menos ( $\beta = 0,92$ ) foram mais assistidos. A maioria dos vídeos utilizaram 6 ou menos caracteres em caixa-alta em seus títulos, mas os que continham pelo menos 8 caracteres ( $\beta = 1,02$ ) foram associados a um maior número de visualizações, ainda que representem apenas 20,11% da amostra.

Com relação ao dia da postagem dos vídeos, sábado e domingo foram os dias em que os mediadores dos canais menos postaram. Contudo, os dias mais associados positivamente a uma maior visualização foram sexta ( $\beta = 0,50$ ), sábado ( $\beta = 0,55$ ), domingo ( $\beta = 0,44$ ) e segunda ( $\beta = 0,43$ ); ou seja, no final e no início da semana. Quando analisado em que parte do dia a postagem foi feita, a maioria foi realizada à noite ou à tarde, não obstante a madrugada ( $\beta = 1,18$ ) ter sido o horário associado a mais *views*.

Finalmente, vídeos designados à categoria “música” ( $\beta = 4,66$ ) estiveram associados a um maior número de visualizações, seguido por “ciência” ( $\beta = 1,56$ ), “comédia” ( $\beta = 1,39$ ), “educação” ( $\beta = 1,37$ ) e “*how-to*” ( $\beta = 1,30$ ). No que se refere ao uso da ferramenta de *tags* do YouTube, os vídeos nos quais os mediadores usaram pelo menos 11 *tags* ( $\beta = 1,66$ ) foram mais assistidos.

Das variáveis ligadas à conteúdo textual dos títulos dos vídeos, conforme descrito na Tabela 4, observam-se associação positiva acerca da presença de elementos ligados à desinformação/controvérsia com relação às vacinas ( $\beta = 1,23$ ) e um maior número de visualizações, assim como o uso de termos associados aos agentes etiológicos ( $\beta = 1,08$ ), aos eventos pós-vacinação ( $\beta = 0,13$  para aqueles frequentes,  $\beta = 0,63$  para os graves), às doenças ( $\beta = 0,89$ ) e aos produtos ( $\beta = 0,24$ ).

**Tabela 4 – Resultados do modelo para variáveis a nível textual dos títulos dos vídeos**

Categoria	Coefficiente	Desvio padrão	p-valor ( $\alpha < 0,05$ )	Intervalo de confiança (95%)	
intercept	8,7165	0,016	0,000	8,685	8,748
doenças	0,8690	0,021	0,000	0,829	0,909
partes_corpo	-0,1645	0,014	0,000	-0,193	-0,136
vacina	-0,0697	0,020	0,001	-0,110	-0,030
produtos	0,2367	0,070	0,001	0,100	0,373
ciência	-0,1878	0,050	0,000	-0,286	-0,090
medicamentos	0,5454	0,119	0,000	0,311	0,779
eventos_freq	0,1765	0,038	0,000	0,102	0,251
eventos_graves	0,6253	0,122	0,000	0,386	0,864
agentes	1,0776	0,039	0,000	1,001	1,154
vac_política	-0,0701	0,071	0,320	-0,208	0,068
controvérsia	1,2327	0,019	0,000	1,195	0,032

Fonte: Elaborada pelos autores.

Ao analisar os resultados da regressão das variáveis referentes às características relacionadas ao mediador dos vídeos, dispostos na Tabela 5, observa-se que o uso de figuras humanas na *thumbnail* dos vídeos ( $\beta = 0,37$ ) foi positivamente associado a um maior número de visualizações. E que essas figuras – de homens ( $\beta = 0,85$ ) e brancos ( $\beta = 0,92$ ) – foram as que mais corroboraram positivamente para um número maior de *views* na plataforma, enquanto as contrapartidas tiveram correlação negativa.

**Tabela 5 – Descrição dos resultados do modelo para as variáveis associadas ao mediador**

Variável	Coefficiente	p-valor ( $\alpha < 0,05$ )	Intervalo de confiança (95%)	
intercept [homem/branca]**	8,2760	0,000	8,114	8,438
intercept [mulher/não branca]**	10,0464	0,000	10,015	10,078
gênero [homem]	0,8475	0,000	0,788	0,907
gênero [mulher]	-0,8475	0,000	-1,079	-0,767
raça/cor [não branca]	-0,9229	0,000	-0,907	-0,788
raça/cor [branca]	0,9229	0,000	0,767	1,079
figura humana***	0,3675	0,000	0,333	0,402

Fonte: Elaborada pelos autores.

\*\* Regressões bivariadas distintas. Apesar de as variáveis raça/cor e gênero serem binárias e correlacionadas internamente, levando à multicolinearidade, os coeficientes foram apresentados para evidenciar os efeitos positivos e negativos na variável preditora.

\*\*\* Regressão univariada.

## DISCUSSÃO

Ainda que não diretamente voltados às vacinas, trabalhos como o de Yang *et al.* (2022) esclarecem alguns dos fatores relacionados a um maior engajamento no YouTube. Analisando o canal Reactions (canal de divulgação científica de química), as autoras indicam uma associação positiva entre títulos mais curtos e concisos e um maior número de visualizações, correspondendo aos achados da presente pesquisa. A hipótese apontada pelas autoras foi a de que esse resultado pode ser consequência do fato de que, quando

títulos mais longos ficam “cortados” na plataforma, ou seja, quando não são apresentados na íntegra aos usuários, não traduzem, portanto, seus conteúdos.

Por outro lado, Yang e colaboradoras (2022) apontam que a duração dos vídeos teve coeficiente negativo com relação ao número de *views*, e, por essa razão, os vídeos mais curtos foram aqueles mais assistidos. Esse resultado também foi encontrado por Munaro *et al.* (2021), mas não coaduna com os achados desta pesquisa, uma vez que os vídeos de média e longa duração apresentaram coeficientes consideravelmente maiores do que aqueles referentes aos vídeos curtos.

A explicação para essa divergência pode estar relacionada aos temas abordados nos vídeos analisados, assim como os momentos sócio-históricos de produção e publicação. A hipótese aqui sugerida é a de que para os vídeos sobre vacina publicados no âmbito da covid-19, os resultados indicam que as pessoas que buscaram informações sobre imunizantes no YouTube tendem a privilegiar vídeos com maior potencial explicativo, diferentemente dos resultados encontrados por Munaro e colaboradores no trabalho supramencionado.

Diante de um evento sanitário da magnitude da pandemia de covid-19, marcada por incertezas de toda a sorte (Lima *et al.*, 2020) – pela politização generalizada, sobretudo acerca das vacinas no Brasil, e pela desordem informacional (Recuero *et al.*, 2021) –, sugere-se que a percepção do risco tenha interpelado o comportamento de consumo de informações na plataforma, associado ao fato de pesquisas apontarem que o sistema de recomendações do YouTube privilegia vídeos longos em detrimento dos de menor duração (Calvo; Cano-Orón; Llorca-Abad, 2022).

Essas circunstâncias são apontadas aqui como hipóteses para que, na mesma medida em que se buscaram vídeos que supostamente pudessem “passar” mais informações, também se observe uma maior inclinação à visualização de vídeos que continham algum elemento associado à desinformação ou à controvérsia em seus títulos, além da menção a termos relacionados aos eventos pós-vacinação, fossem eles graves (trombose, anafilaxia, miocardite, pericardite, trombocitopenia etc.) ou frequentes (coceira, febre, tosse, cansaço, fadiga); ou a termos relacionados às doenças (sobretudo a “covid”), aos agentes etiológicos (especialmente “vírus”) e a medicamentos (hidroxicloroquina, drogas, tratamento precoce, azitromicina etc.). Observa-se também um maior número de caracteres em caixa-alta – já associados à desinformação (Damstra *et al.*, 2021).

Apesar de esses resultados implicarem uma realidade que se demonstra propícia para a desordem informacional, o fato de vídeos inseridos na categoria “ciência” estarem associados a um maior número de visualizações indica que esses elementos foram também usados por aqueles que objetivavam praticar a desinformação, como divulgadores científicos e *youtubers* não especializados.

Com efeito, isso se confirma ao se analisarem os quatro vídeos com maior número de visualizações entre os 13.619 que compõem o *corpus* de análise desta investigação: “Entrevista com Bolsonaro: traição, eleições, vacina, economia e muito mais”; “VACINA COVID: O QUE NÃO TE CONTARAM SOBRE A TERCEIRA DOSE”; “A vacina contra a Covid-19 é perigosa? Anthony Wong responde”; e “COVID: VACINA DE OXFORD ASTRAZENECA CAUSA TROMBOSE???”. Aqueles em caixa-alta pertencem ao canal Olá, ciência!, importante veículo de divulgação científica sobre vacinas no YouTube, cuja relevância já foi destacada em outros trabalhos (Massarani; Costa; Brotas, 2020).

No entanto, autores como van der Linden (2022) apontam que a estratégia de enfrentar a desinformação ou tratar de controvérsias reproduzindo a totalidades ou partes do conteúdo que se visa contrapor (*e.g.* enfatizar que a vacina da Oxford não causa trombose em tom sugestivo) pode ter um efeito contrário, ou seja, ao invés de mitigar e suprimir, acaba por fazer manutenção e propagar esses termos/narrativas em rede, além de se inserir também nas problemáticas do sensacionalismo e do *clickbait*. De todo modo, essa estratégia parece conferir aos canais voltados à ciência melhores condições de disputar audiência com aqueles que propagam desinformação.

Acerca dessa competição, ao analisarem as interações entre grupos pró-vacinas e antivacinas no Facebook, Johnson e colaboradores (2020) evidenciam que, apesar de serem minoria, os antivacina demonstraram ser mais persuasivos, pois conseguiram cooptar maior número de usuários indecisos para seu *cluster*. Uma das razões apontadas se deve ao fato de que, quando os conteúdos compartilhados nos diferentes grupos foram analisados, percebeu-se que aqueles pertencentes aos antivacina eram mais plurais, enquanto os produzidos e compartilhados por pró-vacinas acabaram sendo mais monotemáticos.

Outrossim, Munaro *et al.* (2021) apontam associação positiva entre vídeos postados em horário não comercial, e nas segundas, terças ou quintas-feiras, e um maior número de visualizações. De fato, os vídeos publicados na madrugada e nas segundas-feiras influenciaram positivamente o número de *views*. No entanto, segundo nossos resultados, a quinta-feira foi um dos piores dias para se subir vídeos sobre vacinas no âmbito da pandemia de covid-19; sendo a segunda, a sexta, o sábado e o domingo aqueles que apresentaram melhores *performances*.

Assim como concluem Munaro *et al.*, também se sublinha aqui a necessidade de cautela ao levar esse fator em consideração, uma vez que há a indicação de um contexto-dependente. Ao se postar um vídeo no YouTube, os seguidores do canal são notificados acerca da publicação quando essa opção está ativada. A hipótese aqui apontada é a de que quando isso ocorre no fim da semana, quer dizer, na sexta ou no sábado, a probabilidade de esses seguidores estarem disponíveis para assistir ao vídeo mais prontamente é maior do que durante a semana.

Finalmente, no que tange às variáveis ligadas ao mediador/canal, faz-se importante destacar uma limitação: a existência de figura humana na *thumbnail* dos vídeos não implica considerar que elas sejam representações ou fotografias dos(as) donos(as) dos canais, dado que elas podem também ser figuras ilustrativas ou estarem associadas a pessoas convidadas para interlocução na própria produção.

Por essa razão, o fato dessas figuras serem majoritariamente masculinas e brancas, quando presentes, explicitam três possíveis cenários com relação à menor quantidade de mulheres e pessoas não-brancas encontrada nas *thumbnail* dos vídeos aqui analisados: a) foram menos convidadas para falar sobre o tema vacina enquanto autoridade; b) foram menos usadas para representar a população/tema discutido no vídeo, ou c) produziram menos conteúdo sobre vacinas no YouTube. De toda forma, evidencia-se o caráter estrutural do racismo e do sexismo na sociedade brasileira no tocante à produção de vídeos sobre vacinas no YouTube (Gonzales, 1984).

A esse respeito, estudos sobre as implicações de gênero na ciência e na divulgação científica têm apontado a existência de assimetrias tanto na presença quanto na recepção do conteúdo/trabalho produzido por homens e mulheres. Welbourne e Grant (2015), por exemplo, destacam que há maior prevalência de homens que mulheres como autores de vídeos sobre ciência no YouTube.

Dalyot, Rozenblum e Baram-Tsbari (2022), ao analisarem os comentários feitos aos *posts* de divulgação de trabalhos científicos feitos por mulheres e homens, destacam dois fatores que podem estar associados à menor representatividade de mulheres em plataformas de maior exposição a audiências: o risco de desqualificação do trabalho e a recepção desvantajosa, inadequada ou hostil às produções femininas.

## CONSIDERAÇÕES FINAIS

O presente trabalho, ao buscar compreender quais são os fatores associados ao maior número de visualizações de vídeos em língua portuguesa brasileira especificamente voltados às vacinas no YouTube, apresenta contribuições importantes para a literatura da comunicação em saúde.

De início, este estudo oferece evidências acerca do potencial do uso de elementos ligados à desinformação/controvérsia nos títulos dos vídeos para suscitar audiências com base nos resultados advindos de uma amostra de 13.619 vídeos, confirmando as hipóteses apontadas em trabalhos anteriores de diversos autores. Além disso, como discutido anteriormente, canais de divulgação científica estão fazendo uso da estratégia

de usar elementos ligados à desinformação/controvérsia sobre vacinas em seus títulos, que muitas vezes se mostra um meio eficaz para garantir maior número de *views* às produções.

Outrossim, a presente investigação apresenta uma dimensão prática – seus resultados podem dar bases para melhor orientar a adoção de certas variáveis nativas do YouTube com potencial de suscitar maior número de visualizações, quais sejam: tamanho do título, duração do vídeo, uso de caracteres em caixa-alta, uso de termos associados à ciência e aos sintomas, uso de figuras humanas na *thumbnail*, dia da semana e período do dia ideais para publicação de vídeos sobre vacinas.

Contudo, há algumas limitações que se faz necessário destacar. A primeira delas diz respeito à variável dependente, que aqui se restringiu ao número de visualizações, mas que precisa ser ampliada para outras, como o número de curtidas, *dislikes* e comentários. Covolo e colaboradores (2017), por exemplo, encontraram que, apesar de menos vistos, vídeos antivacinas angariaram mais *likes* e compartilhamentos.

Faz-se necessário também investigar melhor o conteúdo mobilizado nos vídeos para se verificar se aquilo observado por Johnson e colaboradores (2020) no Facebook, com relação à composição do conteúdo produzido por grupos pró-vacinas e antivacinas e seu potencial de suscitar interesse ou convencer pessoas indecisas, se aplica também ao YouTube. Isso poderia ser feito através de modelagem de tópicos ou mediante outra técnica de PLN adequada para a análise de grande quantitativo de dados textuais.

Além disso, a interpretação acerca da raça/cor e do gênero acabou sendo abrangente, quando comparada à motivação inicial de incluí-las na investigação. Todavia, essas articulações certamente levantam questões importantes acerca das assimetrias já documentadas na literatura, produzidas e perpetuadas pela atuação do gênero (patriarcado) e da raça/cor (racismo). Elas sublinham também a importância de novos estudos voltados a essa dimensão na comunicação em saúde, e, mais especificamente, nas produções acerca das vacinas no YouTube.

Dessa forma, os achados desta pesquisa corroboram com a constatação de que o ciberespaço está longe de ser um território neutro ou meramente tecnocrático. Tal como Rosalía Winocur (2006) pontua, a internet não se traduz apenas em virtualidades, pois não é fechada em si mesma, não é fechada hermeticamente aos contextos político, sócio-histórico e cultural em que os indivíduos estão circunscritos. A internet é, na verdade, produzida por “sujeitos reais nos espaços concretos de sua vida cotidiana” (Winocur, 2006, p. 554, tradução nossa). Tampouco se trata de uma cultura autorreferenciada, mas, sim, imbricada nas dimensões práticas e simbólicas nas quais esses sujeitos estão situados.

Portanto, sublinha-se a importância de uma análise pautada em uma leitura sociológica compreensiva (Giddens, 1989) que tem por premissa entender o que as pessoas fazem com as TIC frente aos constrangimentos advindos de suas *affordances* funcionais, tendo em vista que a “lógica da plataforma” é a expressão da “lógica do capital” (Souza, 2020).

## REFERÊNCIAS

BROTAS, Antonio Marcos Pereira; COSTA, Marcia Cristina Rocha.; MASSARANI, Luisa. Enquadramentos e desinformação sobre vacina contra covid-19 no YouTube: embaralhamentos entre ciência e negacionismo.

**Mídia & Cotidiano**, Niterói, v. 15, n. 3, p. 73-100, 2021. DOI: <https://doi.org/10.22409/rmc.v15i3.50954>.

Disponível em: <https://periodicos.uff.br/midiaecotidiano/article/view/50954>. Acesso em: 15 dez. 2023.

BUTLER, Judith. **Problemas de gênero: feminismo e subversão da identidade**. 22. ed. Rio de Janeiro: Civilização Brasileira, 2003.

CALVO, Dafne; CANO-ORÓN, Lorena; LLORCA-ABAD, Germán. Covid-19 vaccine disinformation on YouTube: analysis of a viewing network. **Communication & Society**, Pamplona, v. 35, n. 2, p. 223-238, 2022. DOI: <https://doi.org/10.15581/003.35.2.223-238>. Disponível em: <https://revistas.unav.edu/index.php/communication-and-society/article/view/42099>. Acesso em: 15 dez. 2023.

CHAN, Calvin. *et al.* The reliability and quality of YouTube videos as a source of Public Health Information regarding covid-19 vaccination: Cross-sectional study. **JMIR Public Health and Surveillance**, Toronto, v. 7, n. 7, e29942, 2021. DOI: <https://doi.org/10.2196/29942>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/34081599/>. Acesso em: 15 dez. 2023.

COVOLO, Loredana *et al.* What arguments on vaccinations run through YouTube videos in Italy? A content analysis. **Human Vaccines & Immunotherapeutics**, Austin, v. 13, n. 7, p. 1693-1699, 2017. DOI: <https://doi.org/10.1080%2F21645515.2017.1306159>. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/21645515.2017.1306159>. Acesso em: 15 dez. 2023.

DALPOZ, Gabriel Quirino *et al.* Analysis of YouTube® educational videos on prevention of dental caries. **Research, Society and Development**, v. 11, n. 1, p. e26011124693, 2022. DOI: <https://doi.org/10.33448/rsd-v11i1.24693>. Disponível em: <https://rsdjournal.org/index.php/rsd/article/view/24693>. Acesso em: 15 dez. 2023.

DALYOT, Keren; ROZENBLUM, Yael; BARAM-TSABARI, Ayelet. Engagement patterns with female and male scientists on Facebook. **Public Understanding of Science**, Londres, v. 31, n. 7, p. 867-884, 2022. DOI: <https://doi.org/10.1177/09636625221092696>. Disponível em: <https://journals.sagepub.com/doi/full/10.1177/09636625221092696>. Acesso em: 15 dez. 2023.

DAMSTRA, Ayt *et al.* What does fake look like? a review of the literature on intentional deception in the news and on social media. **Journalism Studies**, Londres, v. 22, n. 14, p. 1947-1963, 2021. DOI: <https://doi.org/10.1080/1461670X.2021.1979423>. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/1461670X.2021.1979423>. Acesso em: 15 dez. 2023.

FIALHO, Inês. *et al.* Enfarte agudo do miocárdio no YouTube – Is it all fake news? **Revista Portuguesa de Cardiologia**, v. 40, n. 11, p. 815-825, 1 nov. 2021. DOI: <https://www.revportcardiol.org/pt-enfarte-agudo-do-miocardio-no-articulo-S0870255121001542>. Disponível em: <https://www.revportcardiol.org/pt-enfarte-agudo-do-miocardio-no-articulo-S0870255121001542>. Acesso em: 15 dez. 2023.

GARBIN, Helena Beatriz da Rocha; PEREIRA NETO, André de Faria; GUILAM, Maria Cristina Rodrigues. A internet, o paciente expert e a prática médica: uma análise bibliográfica. **Interface – Comunicação, Saúde, Educação**, v. 12, n. 26, p. 579-588, 2008. DOI: <https://doi.org/10.1590/S1414-32832008000300010>. Disponível em: <https://www.scielo.br/iccse/a/TPC5B5678dnn9YXBFD3KkrK/?lang=pt>. Acesso em: 15 dez. 2023.

GIDDENS, Anthony. **A constituição da sociedade**. São Paulo: Martins Fontes, 1989.

GONZALES, Lélia. Racismo e sexismo na cultura brasileira. **Revista Ciências Sociais Hoje**, São Paulo, v. 2, n. 1, p. 223-243, 1984. Disponível em: [https://edisciplinas.usp.br/pluginfile.php/7395422/mod\\_resource/content/1/GONZALES%2C%20L%2C%A9lia%20-%20Racismo\\_e\\_Sexismo\\_na\\_Cultura\\_Brasileira%20%281%29.pdf](https://edisciplinas.usp.br/pluginfile.php/7395422/mod_resource/content/1/GONZALES%2C%20L%2C%A9lia%20-%20Racismo_e_Sexismo_na_Cultura_Brasileira%20%281%29.pdf). Acesso em: 15 dez. 2023.

JOHNSON, Neil F. *et al.* The online competition between pro- and anti-vaccination views. **Nature**, Londres, v. 582, n. 7811, p. 230-233, 13 maio 2020. DOI: <https://doi.org/10.1038/s41586-020-2281-1>. Disponível em: <https://www.nature.com/articles/s41586-020-2281-1>. Acesso em: 15 dez. 2023.

KÄRKKÄINEN, Kimmo; JOO, Junggseock. FairFace: face attribute dataset for balanced race, gender, and age. **arXiv**, Cornell, 14 ago. 2019. *Preprint*. DOI: <https://doi.org/10.48550/arXiv.1908.04913>. Disponível em: <https://arxiv.org/abs/1908.04913>. Acesso em: 15 dez. 2023.

LI, Heidi Oi-Yee *et al.* YouTube as a source of information on covid-19: a pandemic of misinformation? **BMJ Global Health**, Londres, v. 5, n. 5, p. e002604, 1 maio 2020. DOI: <https://doi.org/10.1136/bmjgh-2020-002604>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/32409327/>. Acesso em: 15 dez. 2023.

LIMA, Clóvis Ricardo Montenegro de *et al.* Emergência de saúde pública global por pandemia de covid-19. **Folha de Rosto: Revista de Biblioteconomia e Ciência da Informação**, v. 6, n. 2, p. 5-21, 2020. DOI: <https://doi.org/10.46902/2020n2p5-21>. Disponível em: <https://periodicos.ufca.edu.br/ojs/index.php/folhaderosto/article/view/490>. Acesso em: 15 dez. 2023.

LOPES, Arthur da Silva; BROTAS, Antonio Marcos Pereira. Echo chambers and vaccines against covid-19 misinformation on Twitter: machine learning and network analysis-based approach. **Research, Society and Development**, Vargem Grande Paulista, v. 12, n. 2, p. e22812240159, 2023. DOI: <https://doi.org/10.33448/rsd-v12i2.40159>. Disponível em: <https://rsdjournal.org/index.php/rsd/article/view/40159>. Acesso em: 15 dez. 2023.

MASSARANI, Luisa; COSTA, Márcia Cristina Rocha.; BROTAS, Antonio Marcos Pereira. A pandemia de covid-19 no YouTube: ciência, entretenimento e negacionismo. **Revista Latinoamericana de Ciencias de la Comunicación**, São Paulo, v. 19, n. 35, p. 2455-256, 2020. Disponível em: <https://www.arca.fiocruz.br/handle/icict/46507>. Acesso em: 15 dez. 2023.

MASSARANI, Luisa *et al.* Vacinas contra a covid-19 e o combate à desinformação na cobertura da Folha de S.Paulo. **Fronteiras – Estudos Midiáticos**, São Leopoldo, v. 23, n. 2, p. 29-43, 2021. DOI: <https://doi.org/10.4013/fem.2021.232.03>. Disponível em: <https://revistas.unisinus.br/index.php/fronteiras/article/view/22592>. Acesso em: 15 dez. 2023.

MCCLUNG, H. J.; MURRAY, R. D.; HEITLINGER, L. A. The Internet as a source for current patient information. **Pediatrics**, Elk Grove Village, v. 101, n. 6, jun. 1998. DOI: <https://doi.org/10.1542/peds.101.6.e2>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/9606244/>. Acesso em: 15 dez. 2023.

MUNARO, Ana Cristina. *et al.* To engage or not engage? The features of video content on YouTube affecting digital consumer engagement. **Journal of Consumer Behaviour**, Londres, v. 20, n. 5, p. 1336-1352, 2021. DOI: <https://doi.org/10.1002/cb.1939>. Disponível em: <https://onlinelibrary.wiley.com/doi/full/10.1002/cb.1939>. Acesso em: 15 dez. 2023.

NEWMAN, Nic. Overview and key findings of the 2022 – Digital News Report. **Reuters Institute for the Study of Journalism**, Londres, 15 jun. 2022. Disponível em: <https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2022/dnr-executive-summary>. Acesso em: 23 nov. 2022.

NOGUEIRA, Oracy. Preconceito racial de marca e preconceito racial de origem: sugestão de um quadro de referência para a interpretação do material sobre relações raciais no Brasil. **Tempo Social**, São Paulo, v. 19, n. 1, p. 287-308, 2007. DOI: <https://doi.org/10.1590/S0103-20702007000100015>. Disponível em: <https://www.scielo.br/j/ts/a/MyPMV9Qph3VrbSNDGvW9PKc/?lang=pt>. Acesso em: 15 dez. 2023.

NUNES, Georgina Helena Lima. Autodeclarações e comissões: responsabilidade procedimental dos/as gestores/as de ações afirmativas. In: DIAS, Gleidson Renato Martins; TAVARES JUNIOR, Paulo Roberto Faber. (org.). **Heteroidentificação e cotas raciais**: dúvidas, metodologias e procedimentos. Canoas: IFRS Campus Canoas, 2018. p. 11-30. Disponível em: [https://www.geledes.org.br/wp-content/uploads/2019/03/Heteroidentificacao\\_livro\\_ed1-2018.pdf](https://www.geledes.org.br/wp-content/uploads/2019/03/Heteroidentificacao_livro_ed1-2018.pdf). Acesso em: 15 dez. 2023.

PENTEADO, Claudio Luis de Camargo *et al.* #Vacinar ou não, eis a questão! As emoções na disputa discursiva sobre a aprovação das vacinas contra a covid-19 no Twitter. **Política & Sociedade**, Florianópolis, v. 20, n. 49, p. 104-133, 2021. DOI: <https://doi.org/10.5007/2175-7984.2021.85145>. Disponível em: <https://periodicos.ufsc.br/index.php/politica/article/view/85145>. Acesso em: 15 dez. 2023.

PEREIRA NETO, André *et al.* O paciente informado e os saberes médicos: um estudo de etnografia virtual em comunidades de doentes no Facebook. **História, Ciências, Saúde – Manguinhos**, Rio de Janeiro, v. 22, supl., p. 1653-1671, 2015. DOI: <https://doi.org/10.1590/S0104-59702015000500007>. Disponível em: <https://www.scielo.br/j/hcsm/a/NMrcHvYypNG3sFQmvYwv4vR/?format=pdf&lang=pt>. Acesso em: 15 dez. 2023.

PIERRI, Francesco. *et al.* Online misinformation is linked to early covid-19 vaccination hesitancy and refusal. **Scientific Reports**, Londres, v. 12, n. 5966, p. 1-7, 26 abr. 2022. DOI: <https://doi.org/10.1038/s41598-022-10070-w>. Disponível em: <https://www.nature.com/articles/s41598-022-10070-w>. Acesso em: 15 dez. 2023.

RECUERO, Raquel *et al.* **Desinformação, mídia social e covid-19 no Brasil**: relatório, resultados e estratégias de combate. Pelotas: Mídia – Grupo de Pesquisa em Mídia, Discurso e Análise de Redes Sociais, 2021. Relatório de pesquisa. Disponível em: <https://wp.ufpel.edu.br/midiars/files/2021/05/Desinformac%CC%A7a%CC%83o-covid-midiars-2021-1.pdf>. Acesso em: 15 dez. 2023.

RECUERO, Raquel; SOARES, Felipe. O discurso desinformativo sobre a cura da covid-19 no Twitter: estudo de caso. **E-Compós**, Brasília, DF, v. 24, p. 1-29, 2021. DOI: <https://doi.org/10.30962/ec.2127>. Disponível em: <https://www.e-compos.org.br/e-compos/article/view/2127>. Acesso em: 15 dez. 2023.

RIEDER, Bernhard. **YouTube Data Tools**. Amsterdã: Digital Methods Initiative, c2023. Disponível em: <https://ytdt.digitalmethods.net/>. Acesso em: 31 maio 2023.

SERENGIL, Sefik Ilkin; OZPINAR, Alper. LightFace: A hybrid deep face recognition framework. *In: INNOVATIONS IN INTELLIGENT SYSTEMS AND APPLICATIONS CONFERENCE*, 15-17 out. 2020, Istanbul. **Proceedings** [...], Istanbul, v. 2020, p. 1-5, 2020. DOI: <http://dx.doi.org/10.1109/ASYU50717.2020.9259802>. Disponível em: <https://ieeexplore.ieee.org/document/9259802>. Acesso em: 15 dez. 2023.

SOUZA, Rafael Bellan Rodrigues de. A comunicação contra-hegemônica no capitalismo digital: limites e contradições. **Liinc em Revista**, Rio de Janeiro, v. 16, n. 1, p. e5133, 2020. DOI: <https://doi.org/10.18617/liinc.v16i1.5133>. Disponível em: <https://revista.ibict.br/liinc/article/view/5133>. Acesso em: 15 dez. 2023.

VAN DER LINDEN, Sander. Misinformation: susceptibility, spread, and interventions to immunize the public. **Nature Medicine**, v. 28, n. 3, p. 460-467, 2022. DOI: <https://doi.org/10.1038/s41591-022-01713-6>. Disponível em: <https://www.nature.com/articles/s41591-022-01713-6>. Acesso em: 15 dez. 2023.

WARDLE, Claire; DERAKHSHAN, Hossein. **Information disorder: toward an interdisciplinary framework for research and policy making**. Estrasburgo: Council of Europe, 2017. Disponível em: <https://edoc.coe.int/en/media/7495-information-disorder-toward-an-interdisciplinary-framework-for-research-and-policy-making.html>. Acesso em: 31 maio 2023.

WELBOURNE, Dustin J.; GRANT, Will J. Science communication on YouTube: Factors that affect channel and video popularity. **Public Understanding of Science**, Londres, v. 25, n. 6, p. 706-718, 2015. DOI: <http://dx.doi.org/10.1177/0963662515572068>. Disponível em: <https://journals.sagepub.com/doi/10.1177/0963662515572068>. Acesso em: 15 dez. 2023.

WORLD HEALTH ORGANIZATION (WHO). **Ten threats to global health in 2019**. Genebra: WHO, 2019. Disponível em: <https://www.who.int/news-room/spotlight/ten-threats-to-global-health-in-2019>. Acesso em: 23 nov. 2022.

WINOCUR, Rosalía. Internet en la vida cotidiana de los jóvenes. **Revista Mexicana de Sociología**, Cidade do México, v. 68, n. 3, p. 551-580, 2006. Disponível em: <https://www.redalyc.org/pdf/321/32112601005.pdf>. Acesso em: 15 dez. 2023.

YANG, Shiyu *et al.* The science of YouTube: What factors influence user engagement with online science videos? **PLoS ONE**, São Francisco, v. 17, n. 5, p. e0267697, 2022. DOI: <https://doi.org/10.1371/journal.pone.0267697>. Disponível em: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0267697>. Acesso em: 15 dez. 2023.