A r t i c l e s

The public conversation about vaccines and vaccination against covid-19 on Twitter: an infodemiological study

Arthur da Silva Lopes[i]
 https://orcid.org/0000-0001-9137-3184

Antonio Marcos Pereira Brotas[ii]
 http://orcid.org/0000-0001-8438-2445

Luisa Massarani[iii]
 http://orcid.org/0000-0002-5710-7242

[i] (Universidade Federal da Bahia, Instituto de Saúde Coletiva, Programa de Pós-Graduação em Saúde Coletiva. Salvador – Ba, Brazil).

[ii] (Fundação Oswaldo Cruz, Instituto Gonçalo Moniz. Salvador - Ba, Brazil).

[iii] (Fundação Oswaldo Cruz, Casa de Oswaldo Cruz; Instituto Nacional de Comunicação Pública da Ciência e Tecnologia. Rio de Janeiro - RJ, Brazil).

Abstract

The present research is infodemiological and aims to characterize the public conversation about the vaccine and vaccination against COVID-19 in Brazil that took place on Twitter. To map the latent topics in the collected data (13.2 million tweets), topic modeling was used. since the effectiveness of public vaccination policies is conditioned by its population adherence. From the results, 50 topics were identified and discussed over the period of analysis, 11 of which have direct potential to promote vaccine hesitancy, categorized into thematic groups, namely Accountability, Public Policies, Conflicts, Bioethics, Experience Reports, Economic Aspects, Conspiracy Theories; and Safety, Efficacy and Importance of Vaccines. There was a strong politicization translated by the most retweeted profiles

and hashtags used, not only partisan, but mostly civic, represented by users' participation through the manifestation of social control in health.

Keywords: Vaccine. Vaccination. COVID-19. Twitter. Infodemiology.

Introduction

The beginning of the COVID-19 pandemic course was overly marked by uncertainty, afterall, little was known about the new coronavirus. However, this scenario of doubt was preceded by a concrete bet: the use of vaccines to effectively control this public health crisis.

In this respect, a global race was launched to develop and acquire vaccines against COVID-19. However, this technological venture has run up againstthe social-cultural dimension inherent in the adhesion to the vaccination plan by the population, which includes the individuals' informational diets and perception of risk.

To make matters worse, the information was contested in the context of the pandemic and became a strategy of power. (LIMA et al., 2020). Thus, there was an accentuation of the informational disorder, characterized by the dissemination of considerable amount of information to the point where it becomes difficult or even impossible to identify what is true and what is false, phenomena named "infodemics" by the the World Health Organization (WHO) (CINELLI et al., 2020).

The emergence and development of the internet has played a fundamental role in shaping infodemics, as it has led to major changes in the production, dissemination and consumption of information. One of the facets of these changes is the arbitrary dissemination of false information on websites and social media platforms. Such content is mostly produced by anti-vaccine movements, which portray vaccines as posing greater risks to human health than the disease they aim to prevent (LIMA et al., 2020).

This issue is not unique to the COVID-19 pandemic, the role of these movements and the decline of vaccination led the WHO, in 2019, to list vaccine hesitancy as one of the ten public health threats to be addressed in the coming years (WHO, 2019).

Vaccine hesitancy is associated with a set of beliefs, attitudes and behaviors including refusal, delay or reluctance to accept an individual's (or their caregiver's) decision to be vaccinated, despite vaccine availability and the risk posed by a particular immunopreventable health problem (KANG, 2017).

In this sense, Social media platforms such as Twitter play an important role in shaping public perceptions about vaccines and vaccination public policies, as they are increasingly becoming the most preferred sources of health information (HAMAN, 2020) in Brazil, only behind television (AVAAZ, 2020).

The main activity carried out on this platform is microblogging, which allows anyone who has something to share to have access to different audiences (ZAGO, 2019). When people

share events and daily life experiences on Twitter, it inevitably becomes a rich place for opinion mining on a variety oftopics, including health.

On the other hand, Recuero and Soares (2021) point out that platforms such as Twitter can also be understood as spaces of dispute. Whether as a microblog, or as a political space of discursive dispute, the information that circulates on these platforms is usually qualified according to emotions, personal beliefs and political convictions; and is evaluated according to its capacity to generate interactions (SINGER, 2013; DEUZE, 2005 apud AMARAL; SANTOS, 2019, p. 65).

In this way, it creates a fertile territory for these contradictory narratives about vaccines to gain relevance, which given the specific nature of the object that is being distorted, becomes particularly worrying, due to the persuasive power of the information that is located within the health issue (MONARI; BERTOLLI FILHO, 2019).

Therefore, infodemiological studies have gained relevance due to the number of circulating contents (big data) and the aforementioned importance of social media platforms in Public Health (ABD-ALZARAZ et al., 2020; ALAGHA, 2021; BOON-ITT; SKUNKAN, 2020; MELO; FIGUEIREDO, 2021; YIN et al., 2022; YOUSEFINAGHANI et al., 2021).

Therefore, we aim to conduct an infodemiological study aiming to understanding the public debate about the development of vaccines against the new coronavirus and its application in terms of public policy in Brazil, using Twitter as a site of analysis.

In terms of questions of analytical interest (RQ), there are: RQ1: What are the general metrics of the conversation (number of tweets and use of hashtags and mentions)? RQ2: Which users participated in the conversation? What is the relative influence that these actors and actresses had in shaping the debate on the platform? RQ3: What are the trend lines of the conversation activity over the time period covered by the analysis? RQ4: What topics are users covering in the conversation about vaccines and vaccination policy on Twitter?

Methods

The methodology is semi-automated and divided in: (1) data collection, (2) preprocessing, (3) topic modeling and (4) analysis.

Data collection

In this work, we chose to use Netlytic platform (VIS, 2013) due to the practicality conferred by previous experience and extensive documentation related to its use. In this sense, the data collection was made based on the descriptor "vaccine OR vaccination", from December 11th, 2020 to May 2021. This period coincides with the repercussions of the events related to the formulation vaccination program, the emergency authorization of the vaccines then available and the effective operationalization of the vaccination in different countries and regions, such as the United Kingdom on December 2nd, 2020 (BBC BRASIL, 2020).

Intercom, Rev. Bras. Ciênc. Comun., São Paulo, v. 46, e2023121, 2023

3/21

With the data collection being initiated in December, we would cover a part of the debate about vaccination in the field of COVID-19 pandemic. It is important to mention that the collection was done automatically, in a way that the datasets were created from the continuous extractions of tweets every 15 minutes, until they reached 100 thousand tweets. After reaching this amount of data, another dataset (tabular structures similar to a spreadsheet) was then created and the same process was repeated until the last day of data collection.

Preprocessing

At this stage, the datasets were merged into one. To identify the hashtags used by the users as well as the users that were mentioned in the tweets we performed an extraction process of the words associated with "#" and "@" respectively.

In addition, user profiles were identified based on the "author" and "user_bio" metadata, which consists of the user's name and the biography that they wrote on Twitter, through an active search of regular expressions related to the profiles of analytical interest (see two examples in Table 1).

Table 1 - Examples of the profiles' categories and terms used to classify the users

Category	Terms**
Politician	Name of the politician, r'deputy', r'senator', r'councillor', r'mayor', r'minister'.
Science communicator	*r'\bmsc', '\bphd', r'phd.*in', "epidemiologist", 'science communicator',"science communication", "scientist", "sanitarist"*

Source: elaborated by the authors.
**In *regex* notation

The categories used to classify the profiles of the participants in the conversation were defined after an exploratory analysis, from which the following were identified: politician, political profile, science communicator, journalist, organization, influencer, health professional, news outlet, partisan media, artist and citizen.

In order to preserve anonymity and to comply with Twitter's privacy policy, users who are not public figures were not mentioned individually, and all those whose profile could not be classified were grouped together in the category "citizen".

Topic modeling

To answer the question "What are the main topics discussed over the conversation?", topic modeling was used. Among the set of algorithms belonging to unsupervised machine learning that perform this task, the BERTopic was chosen due to the fact that it presents better results overall (ABUZAYED; AL-KHALIFA, 2021; SILVEIRA et al., 2021).

As a consequence of the sample size (tweets-replied + original posts) being equivalent to a total of 2.8 million tweets, a stratified sampling was done in the basis of the publication month (1% of the tweets) so that the samples from each posting day could be captured throughout the whole period, while maintaining the proportion of tweets monthly. Thus, the subsample has a total of 144,026 tweets, with which the topic modeling step was performed.

Results and discussion

The methodology operationalization resulted in a corpus of just over 13,3 million tweets. However, tweets with the same ID number, authorship, publication date and textual content were considered duplicates and, consequently, removed. Finally, the analyzed sample consisted of 13,2 million tweets, of which 7,2 million were retweets (67%); 1,4 million werereplies (13,6%); 1,4 million wereoriginal tweets (13,6%), and 637,488 were quotes (5,8%).

Timeline of the conversation

Based on the activity of the conversation shown in Figure 1, it was possible to identify 03 stages: (a) from December 11, 2020 to February 2, 2021; (b) from February 23 to April 5, 2021; and (c) May 31, 2021. As far as thegeneral trends are concerned, there were peaks that stood out during the analysis period, showing that the intensity of the activity was not homogeneous, although it remained in a certain pattern within each phase.

Figure 1 - Timeline of the conversation activity about vaccines on Twitter
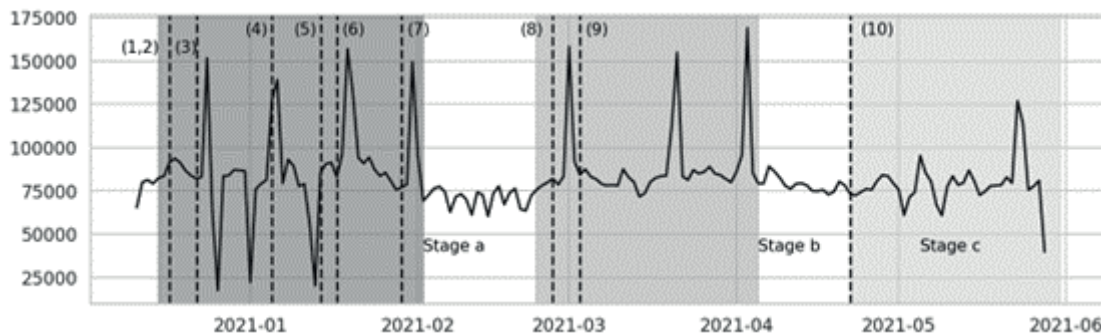
Legend: expression "(number)" is being used to represent the key events that occurred in the context of COVD-19 pandemic in Brazil
Source: elaborated by the authors.

Table 2 - List of the events responsible for causing peaks during the conversation, dates of occurrence and correspondent peaks

Id	Event	Date	Peak
1	STF decides for compulsory vaccination	17/12/2020	24 dec
2	Bolsonaro called into question vaccines safety and efficacy: "You will become an alligator".	17/12/2020	24 dec
3	Supreme Court asks Fiocruz to reserve vaccines for public servants	22/12/2020	5,6 jan
4	"Brazil is broken. There is nothing that I can do", Jair Bolsonaro	05/01/2021	5,6 jan
5	Crisis in Manaus	14/01/2021	20 jan
6	Beginning of vaccination against COVID-19 in Brazil	17/01/2021	19 jan
7	Bolsonaro's government releases 3 billion reais to parliamentarians amid election period in the Congress	29/01/2021	31 jan
8	Return of the restrictive measures in brazilian cities	26/02/2021	1,21 mar
9	"Brazil's Covid Crisis Is a Warning to the Whole World", The New York Times	03/03/2021	21 mar, 3 abr
10	Beginning of the COVID-19 Parliamentary Committee of Inquiry (PCI)	22/04/2021	23,27,28 mai

Source: elaborated by the authors.

As can be seen in Figure 1 and Table 2, thehighlighted events served as discussion triggers, in some moments, but all, except the events 4 and 10, had the same half-life of two days or more. In addition, the activity peaks were not conditioned to the days on which they occurred. Even when a discussion was triggered, the corresponding peak did not necessarily translate a discussion strictly related to the event in question.

For example, the highest peak in the timeline, on April 3rd, 2021, was the most retweeted post by Atila Lamarino overall, as it triggered more retweetactivity that day: "The government: \n Does not test \n Does not provide assistance \b Dismisses vaccines \n Promotes agglomeration \n Does not make lockdown as public policy and says that it doesn't work \n\b 'Ah, but the COVID is like that because Brazilian won't buckle down'".

However, other topics were present, such as: the reduction in the number of deaths in the United Kingdom, from 1,200 to less than 20; the president of Argentina, Alberto Fernandez, who tested positive for COVID-19 even after being immunized with Sputnik V; the vaccination of the then ex-president "Lula"; and the deficit of people who did not complete their vaccination schedule because they did not want to take the second dose of vaccines.

6/21

Intercom, Rev. Bras. Ciênc. Comun.,
São Paulo, v. 46, e2023121, 2023

General metrics of the conversation

Overall, the tweets were composed of 54,2% of devices beyond the text itself, being: 21,4% of mentions to other profiles, 20,1% of URL and 12,7% of hashtags. Although the links are the second most used resource, they are not further described here due to access restrictions to their actual domains.

As for the hashtags used in the conversation, it was shown in Table 3 the 5 most relevant. Just by reading them it is possible to infer the character of the debate, where it can be seen that most are related to the mobilization with relation for the availability of vaccines for everyone (#VaccinesForEveryone, #ArmyForVaccines and #AllForVaccines), making a total of 14,767 uses. In addition, those associated with holding the then president of Brazil accountable stood out (#BolsonaroOut and #BolsonaroGenocidal), having been used 10,921 times.

Table 3 - List of the five most used hashtags in the conversation

	Hashtag	Total
1	#OutBolsonaro	6,448
2	#VaccineForEveryone	5,703
3	#AllForVaccines	4,869
4	#BolsonaroGenocidal	4,473
5	#ArmyForVaccines	4,195

Source: elaborated by the authors.

The relative influence of the most relevant profiles within the conversation

The relative influence of the mapped profiles in the conversation was analyzed based on the number of retweets that the users were able to generate, because among the values that a retweet can assume, there are: referentiality and access. The first refers to the ability of the retweeting activity to give visibility to the message being shared and credit for the tweet content to both the author of the original post and those that retweeted it, which can be translated as influence (RECUERO; ZAGO, 2012).

The second concerns the process of gatekeeping that users perform on the platform, since "when someone RTs, they are passing on information that they believe their network has not yet received. They are therefore providing access to something you consider relevant." (ibdem, p. 24, free translation).

Intercom, Rev. Bras. Ciênc. Comun.,
São Paulo, v. 46, e2023121, 2023

7/21

However, identifying a user's influence based on retweets requires caution, because there are two possible outcomes for this calculation: its nominal value and its real value. The first refers to the number of times a user has been mentioned with the retweet tag ("RT @user"), but a cold look at this data could lead to partial interpretations regarding their actual influence in the conversation.

This is due to the fact that a user can retweet the same post multiple times, increasing its nominal value without adding retweets from new users. In this sense, the real value is related to the number of times that a tweet has been reproduced by unique users (UU). Thus, what can be seen in Table 4 thereby is that Bolsonaro was the one who had more retweets, but most of his content was shared by the same users (only 27% of UU).

Table 4- List of the users most retweeted

Nome	Perfil	Total RT nominal	Total RT real	Total usuários únicos*	Total publicações*	Proporção real/ nominal (%)	Proporção UU/ nominal (%)
@jairbolsonaro	Politician	190,404	132,937	51,599	871,938,756	69.8	27.1
@oatila	Science communicator	136,811	103,834	65,935	142,008,279	75.9	**48.2**
@minsaude	Organization	31,982	22,183	11,468	324,387,049	69.4	35.9
@jdoriajr	Politician	38,569	26,500	15,033	359,680,596	68.7	39.0
@g1	News outlet	55,760	41,546	28,378	923,967,854	74.5	50.1
@butantanoficial	Organization	44,048	33,946	24,694	693,509,105	**77.1**	56.1
@folha	News outlet	38,994	28,005	20,574	648,984,396	71.8	52.8
@dadourado	Science communicator	44,796	31,704	21,384	622,543,584	70.8	47.7
@leandroruschel	Politician	42,004	25,883	15,807	380,491,468	61.6	37.6
@carmelonetobr	Politician	41,161	26,186	13,067	357,607,710	63.6	31.7

*Unique users among those that reproduced the content of the profile in question (real RT).
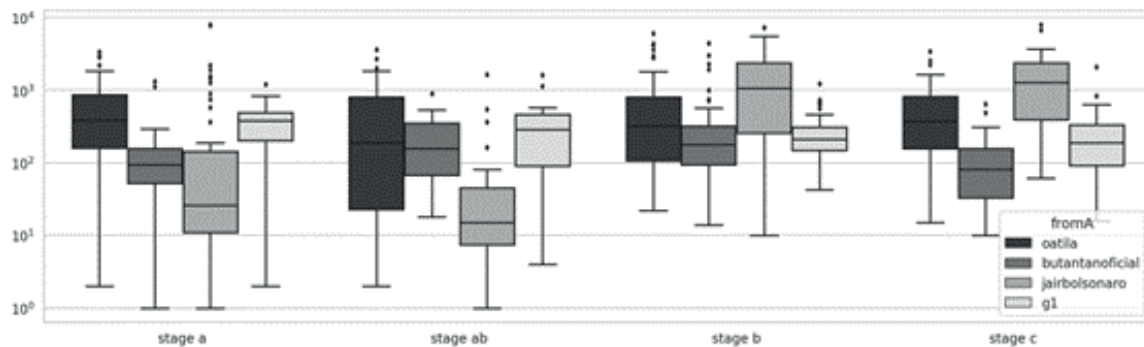Source: elaborated by the authors.

Conversely, Atila Lamarino, the second in total number of retweets, was much more effective in getting different users (48%) to share his content. Moreover, political profiles, politicians and the Ministry of Health (under Bolsonaro's government) were the ones with the lowest proportion between real and nominal retweets.

It is important to note that quotes and retweets of retweets are also tagged with "RT". To quote means to retweet a post adding a comment to it, while the latter refers to users who have retweeted a content from a given profile, and, with that retweet already in their social network, have made it be shared by other users.

For this reason, the nominal value of a retweet cannot be interpreted directly as agreement with the author of the original tweet, because, through quotes, it is also possible to contest the message in question, as it was observed in the case of Jair Bolsonaro, Ministry of Health, and João Dória. Furthermore, the question arises as to whether, in terms of the actual values of retweets, there was a variation in influence depending on the period of analysis.

In fact, as can be seen in Figure 2, at the stages a (December 2020 and January 2021) and b (February 2021), Atila Lamarino was the most influential person in the conversation - although the number of retweets decreased between the periods -, while Jair Bolsonaro was less influential.

Figure 2 - Graph indicating the variation in the number of retweets according to the temporal stage of the conversation

Source: elaborated by the authors.

However, in stage b, the discourse produced by Bolsonaro began to interpellate more people, who began to share his content in a way that allowed him to surpass the other profiles in terms of influence. At this point, the interpretation of the results indicates that there was a change in the tone used in the posts made by Bolsonaro throughout the conversation (Figure 3), which was also demonstrated by Monari et al. (2021).

Figure 3 - Word clouds from Jair Bolsonaro's tweets throughout the conversation

(a) December 2020   (b) From January to May 2021

Legend: most frequent words (a) - "having certification", "official certification", "legal principles", "scientific guidelines", "efficacy", "registry", "obligatory", "meeting resources", "will make available". Words (b) - "million doses", "thousand doses", "emergency aid", "lots of vaccines", "delivered million", "President ", "doses distribution", "employment".
Source: elaborated by the authors.

Initially, the discourses were constructed to defend a position against immunizers under the aegis of the need for Anvisa certification in the face of the supposed unsafe nature of vaccines, as well as mobilization against mandatory vaccination (Figura 3a). But, these began to represent, as aforementioned authors also pointed out, a dispute for symbolic capital between Bolsonaro and the then governor João Doria, over who would be the one considered responsible for the first vaccination in the country. Or, ultimately, over who would be the most competent in regards to the vaccination of the population.

Stage b, the one in which these new narratives become more relevant, is the moment considered the worst of the pandemic, with the eminent/effective collapse of the complementary and public health systems in different states of Brazil, which led to the return of restrictive measures and the consequent polarization between the federal, state and municipal government spheres regarding how to manage the pandemics. For this reason, the discussion revolved around the following question: who would be the one to blame for the situation that Brazil was facing in this moment of the pandemics?

Although the COVID-19 PCI began at the end of April, the request for its establishment was filed on January 18th, 2021. Therefore, agents of the federal government, of the Brazilian fair right - and their supporters -, began to share actions to confront the health crises regarding the acquisition of vaccine doses and application of immunization in the population so that a counter-discourse was to be constructed: "How could the president be the one held responsible for the all the omission if there are multiple actions undertaken by him shared through social media"?

In terms of impact, this seems to have been an effective strategy (see stages b and c) - including to co-optation of new users -; although it competed with great fact-checking mobilization,

through which Jair Bolsonaro was pointed to be oversizing the number of doses and people immunized, and distorting information by making comparisons between the immunization status of different countries (to highlight the supposedly success of his vaccination program).

When we extend the analysis of influence to the profile categories, we see the presence of citizens (92,567, 90%), political profiles (4.252, 4.12%), journalists (2.534, 2,52%), health professionals (1.173, 1.13%), politicians (875, 0.84%), news agencies (626, 0.60%), science communicators (458, 0.44%), organizations (211, 0.20%), digital influencers (146, 0.14%), partisan media (64, 0.06%) and artists (51, 0.05%).

The variation in their influences is shown in Figure 4, where science communicators, news agencies, journalists and politicians, in particular, argue over relevance. However, politicians won this confrontation at all stages, which is a worrying scenario given the role that political leaders play in the profusion of contents capable of arousing, or deepening, the polarization around vaccines (PASCHOALOTTO et al., 2021).

Figure 4 - Graph indicating the variation in the number of retweets according to the temporal stage of the conversation by category, except citizens.
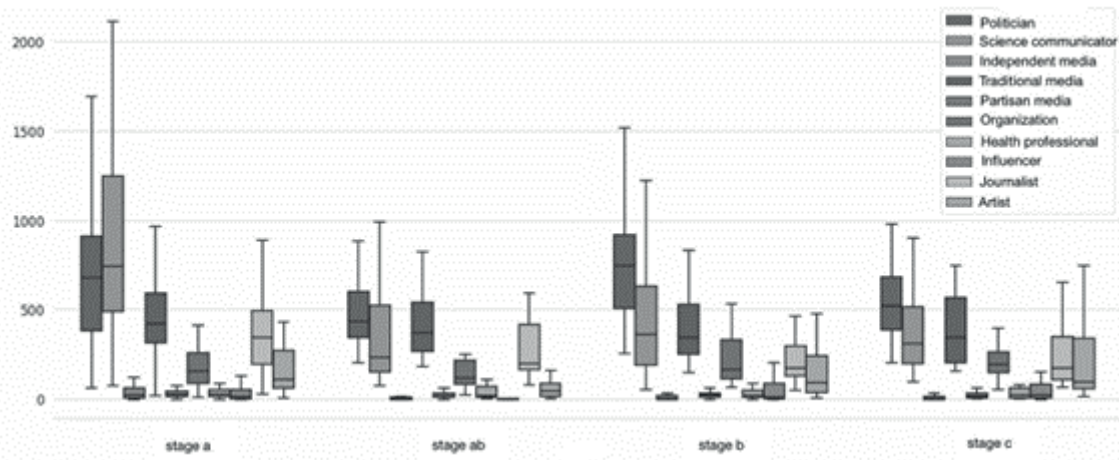
Source: elaborated by the authors.

Which themes were raised over the conversation?

Through topic modeling it was possible to identify 481 topics within the sample given to the algorithm as input, which were later reduced by BERTopic's own mechanisms to 51 (aggregation of those that were too similar) so that a better reading and interpretation of the topics found could be reached.

Originally, the "name" of the topic, i.e., the relationship between the word that describes it – in ascending order of importance -, is presented as: "topic's number_term 1_term 2_ term

Intercom, Rev. Bras. Ciênc. Comun.,
São Paulo, v. 46, e2023121, 2023

11/21

3_term 4". However, since the first term is the most relevant, and for a better presentation of the results, it will be the only one used to illustrate the topics. Thematic categorization was done through extensive reading of the tweets attributed to each topic, and the result is shown in Table 5.

Tabela 5 - List of the themes and its respective topics - name and total number of documents attributed.

Theme	Topics	Name
Accountability	5,6,7,41,42,26, 27, 2, 40	5_genocide, 6_impeachment, 7_israel, 41_ignorou, 42_condensed milk, 26_anvisa, 27_oxygen, 2_china, 40_diplomacy, 16_economy, 23_unemployment
Public Policies	0, 11, 39, 47, 21, 3, 2, 40	0_dose, 11_million, 39_temperature, 47_export, 21_football, 3_teachers, 2_china, 40_diplomacy, 26_anvisa, 5_genocide
Personal Experiences	17, 22, 28, 30, 33, 35, 36, 48	17_dream, 22_waiting, 28_marcos pontes, 30_sleep, 33_medicalrepport, 35_bar, 36_jealousy, 48_gay, 19_carnival
Conflicts	50, 24.2	50_communist, 24_never against, 2_china
Vaccine Effectiveness	1, 4, 8	1_coronavac, 4_sputnik, 8_mutation_virus, 49_baby_milk
Bioethics	21, 18, 25, 29	21_football, 3_teachers, 18_freedom, 25_racism, 29_indigenous
Economic Aspects	16, 23, 31, 40, 2, 47	16_economy, 23_unemployment, 31_tourism, 9_patents, 40_diplomacy, 2_china, 47_exporting
Vaccine Safety	12, 32, 38	12_pregnant, 32_thrombosis, 38_allergic reaction, 49_baby_milk
Vaccine Importance	46	46_polio
Conspiracy Theories	13, 34, 37, 51	13_fake news, 34_whatsapp, 37_bill_gates, 51_facebook
Others	44, 45	44_dengue, 45_cancer

Source: elaborated by the authors.

Personal experiences

The Personal Experiences topic refers to the users' own personal testimonies or other parties regarding their therapeutic journeys, experiences, preferences, expectations, frustrations and hopes related to vaccines and vaccination policies.

The mostretweeted tweets were: "i love you old lady getting ready for her vaccine i love you", "I never imagined that in 2021 we would have to defend vaccines and books", "me when it's my turn to get vaccinated", and "My summer vacation after the vaccine" representing the frequent associations between vaccinations and the possibility of returning to life prior to the pandemics.

Moreover, during the activity of sharing tweets about finally getting vaccinated, the users frequently mentioned the Brazilian United Health System (SUS) and science, giving them credit for the immunization in the country through expressions such as "Long live the SUS! Long live science!", and the hashtags namely #LongLiveSUS (1,221 times) and #DefendTheSUS (2,149 times), #BB21 (5,955 times), #vaccine (2,976 times), #VaccineYes (2,888 times) and #LulaVaccinated (2,449 times).

Public policies

This theme concerns the news and debates in relation to the government actions, whether of its agents or its institutions, aiming to mitigate the COVID-19 pandemic effects. In this sense, tweets related to health surveillance (eg. epidemiological bulletin), to vaccination policy, to restrictive measures and to initiatives of purchasing supplies and vaccines by the executive or legislative branch were part of the discussions.

Among the most relevant tweetsrelated to this topic, "argentina: getting vaccinated and legalizing abortion bra: zil" was the one with the most retweets, emphasizing that there was no vaccination program in the country at that time. Besides that, the one by Dr. Ethel Maciel about the false signatures on the vaccination plan sent to the Superior Federal Court by the Ministry of Health was also considerably retweeted.

In addition, messages with quantitative data on the number of people vaccinated in Brazil and its relative position in relation to the global ranking of the immunization status of each country were frequent, usually associating the Brazilian supposed success with the federal government, particularly with Jair Bolsonaro, as already discussed. For exemple, there is a tweet in which it is affirmed that

> "We have passed the mark of 50 MILLION covid vaccines made available. In total, our government has already distributed 53.4 million doses to the states, of which 28.4 million have been applied. We are the 5th most vaccinated country in the world. With all of them applied, we would be 4th."

Intercom, Rev. Bras. Ciênc. Comun., São Paulo, v. 46, e2023121, 2023

13/21

Accountability

In what concerns the theme of Accountability, the issues discussed can be translated as a manifestation of social control in health, even if they are represented in media spaces through individualized collective political action (AMARAL; SANTOS, 2019, p. 66).

In this sense, accountability represents a digital public mobilization related to holding public agents and government institutions responsible for their actions and omissions regarding the response given to the pandemic. Specifically, the explicit refusals, the omissions in the acquisition of supplies and vaccines by the federal government, and the statements made by politicians, especially Jair Bolsonaro (the most common hashtag was #OutBolsonaro); but also Joao Doria and Eduardo Pazuello, former Ministry of Health referred to as "General" or "the one who understands logistics".

As examples, there are: "Anvisa's criminal sabotage of the sputnik vaccine. they will have to answer in the genocide PCI", and "Anvisa has approved the definitive registration of the pfizer vaccine. yes, the one with 94% efficacy and that bolsonaro didn't want to buy in august 2020, when we were offered 70 million doses".

Economic aspects

Theme where the economic aspects related to the vaccination program were addressed. Topics included the purchase of immunizers by the private sector and the marketing dimension implicated in the process of vaccine acquisition, either from negotiations between countries - mainly China (topic 2) - or with pharmaceutical companies. In this context, a user points out that:

> [...] China is the supplier of vaccines to the whole world, we don't want to run out of supplies and vaccines because the shitty president and his followers keep attacking China, causing diplomatic problems and hindering the purchase of supplies and vaccines.

Additionally, patent breaking (topic 9), vaccine prices and the mechanism of importing and exporting supplies and vaccines (topic 47) - including market reserve, shortage of immunizers and geopolitical asymmetries in vaccine distribution -, were also discussed. In response to the opposition to breaking the patent, users emphasized, in short, that "[...] money/capital takes precedence over health and life. This is what we see, [...]".

Bioethics

It presents the debates of vaccination, highlighting the ethical limits of compulsory vaccination, individual freedom, and the attitudes of health professionals, especially nurses, in

the relation to vaccination. There was also a debate about the prioritization of certain individuals in relation to the criteria of selecting those who should be prioritized in Brazil's vaccination policy guidelines.

Some users emphasized the necropolitical dimension of this operation (BENTO, 2018), as also pointed out by Souza and Santos (2023) when analyzing the case of the population in Brazilian prisons. In addition, there was a movement to denounce fraud, clandestinity and "scams", frequently portrayed as "skip-the-line".

Among the topics that promoted greater engagement, there is the case of businessmen vaccinated with a clandestine vaccine administered by a fake nurse in the city of Belo Horizonte, which the Federal Police later discovered was a saline solution; cases in which nurses simulated the administration of vaccines, referred to as "wind vaccines"; compulsory vaccination and the sanctions due to unethical professional practice.

They were represented by tweets such as: "vaccinating workers who have to leave the house to put food on the table: vaccinating first-year medical students who go out partying every week and are supported by their parents:", "Authoritarianism! MPT says that those who don't take the vaccine could be fired for just cause. the coercive state disrespecting the constitution and infringing one of the fundamental rights of citizens: freedom of choice.", and "professional who administered covid-19 vaccine to elderly women apparently without injecting the liquid and removed from the scale, says g1 secretariat".

Conflicts

Refers to the discussion in which vaccination is framed as a political strategy, and the actions undertaken by institutions and public agents as arising from personal or political interests. Besides, there is also a debate about the discursive disputes engendered by political-ideological confrontations, but also around individual's sociocultural references and their tensions with the system of epistemic authority as addressed by Lavazza e Farina (2020).

Thus, when approaching vaccines in conjunction with governor Joao Doria, the narratives are commonly associated with risks and uncertainties, due to the almost plasma binding between him and the CoronaVac vaccine, which, as an expression of these tensions, comes to be called: "Doria's vaccine", "vachina", "chinese vaccine" and "xing ling vaccine"; the coronavirus as: "chinese virus", "thing from China"; and Doria as: "Doriana", "tight panties", "Agripino", "bold pants"; therefore xenophobic, misogynistic and homophobic speeches.

To illustrates, there are the tweets: "there goes Doria, with tight panties and everything, trying to surf in the vaccination wave. If it had been Bolsonaro holding a public event to give the first vaccine, what would the press have called him? but since it was Doria, the chosen one of Chinese capital, our press would have called him "the first vaccine".

Intercom, Rev. Bras. Ciênc. Comun., São Paulo, v. 46, e2023121, 2023

15/21

Topics associated with vaccine hesitancy

In addition to the themes presented, topics (1, 4, 18, 8, 12, 32, 38, 13, 34, 37 e 51) in which there was greater potential to promote hesitant feelings about vaccines, were also identified, raising doubts about their safety, importance and efficacy, factors already listed in the literature as key elements capable of undermining the population's confidence (FIGUEIREDO et al., 2020).

In terms of the conversation about Vaccines Effectiveness, the CoronaVac vaccine was the one that was the most subject of hesitant discourse about vaccination. In this topic, users often constructed their tweets based on quantitative data, referring to vaccines by their percentage of effectiveness.

For example, a user asks the following question: "effectiveness of vaccines: pfizer 95%, moderna 94.5%, sputinik v 91.4%, oxford 70%, covaxin 63%, coronavac 50.38%, which one will you take?". That is, just like this one, not only tweets with stances against vaccination were posted, but also against vaccines that based on the individuals' own evaluation do not meet their criteria of cost/benefit.

Cost, because those narratives occur concomitantly with those that frame the vaccines as a risk, as in: "desperation running wild...according to scientists, the Chinese vaccine is only 50% effective, i.e. it doesn't protect against the virus at all. apart from the side effects that only god knows, [...]".

Regarding Vaccine Safety, this theme was made up of the topics 12_pregnant_women, 32_thrombosis_risk_astrazeneca_events and 38_allergic reaction_llergic_reaction. It includes tweets characterized by doubts, satire, and reports of adverse effects.

With regard to the former, there were doubts about the safety of vaccines for specific population groups, such as pregnant women, the elderly and children, which are legitimate given the scenario of uncertainty and infodemics. But, also questions that were only intended to incite fear and hesitation.

In this sense, there were tweets pointing out alleged adverse effects of the vaccines used in the brazilian vaccination policy - mainly CoronaVac, Pfizer and Astrazeneca/Oxford - by sharing personal experiences or those of other people (family or friends), in which they reported allergic reactions, encephalomyelitis, myocarditis, thrombosis, blood clots, hemorrhage, inflammation or "heart attacks".

In this vein, there were also tweets suggesting that vaccines caused people to become infected and sick with COVID-19 or other diseases, such as AIDS. In addition to these discourses, vaccines were often associated with death, suggesting that they were the cause of death: "almost two hundred Americans died after taking pfizer's vaccine within two weeks", and "the first indigenous person to take coronavac in Acre dies. unfortunately, we will be guinea pigs for this vaccine...".

Regarding the specific theme of Conspiracy Theories, it suggested that the CoronaVac vaccine would have a microchip in its composition to be implanted in the population given that China or Joao Doria would have a plan to control and monitor the people. In addition to these

theories, users affirmed that the vaccines would provoke genetic mutations, or even that they would be "marks of the beast"; and for those reasons, people should not get vaccinated.

On the other hand, in the midst of the discursive disputes engendered by politicization, there are also counter-narratives to the disinformative messages, as those present in the topics 13, 34, 37 and 51.

Satire was a discursive modality considerably used in conversation to counter these hesitant narratives, as in: "'I don't believe in the efficacy of the vaccine'. my friend, you believed that the cruise was going up.", or in "'for me to get vaccinated, the vaccine has to be at least 90% effective' what kind of requirement is that, I studied with you, you studied to pass with 5".

When addressing the Vaccines Importance (IV), users mention it in the milieu of other COVID-19 control measures, discursive strategy that is present in the specific theme of IV. In addition to this one, the vaccines are framed as an important prevention tool, and as such, are responsible for saving lives (topic 46).

Final remarks

Considering overall findings of the present research, they are in line with what was found in previous studies (OLIVEIRA et al., 2021; RECUERO et al., 2020). In other words, there was a strong politicization in the midst of the COVID-19 pandemic in Brazil, expressed through the higher number of political agents among the most retweeted users in the analyzed sample.

However, what is suggested here is that the polarization is not only in the political-ideological field, but also due to public mobilization of a civic nature, expressing the interstice between media practices and civic involvement (DAHLGREN; ALVARES, 2014). This happened insofar as Twitter became a means through which citizens debated about the vaccination program development, evaluation and monitoring, among themselves and directly with institutions and public agents, using Accountability as a protest platform.

Furthermore, the results also suggest that although Twitter has a real-time publishing character, making it a propitious environment for the easy dissemination of new information, its impact does not imply punctual consequences, because it remains active as topics of conversation over long periods of time.

Regarding the influence of the profiles that participated in the conversation about vaccines and vaccination policies on Twitter, Jair Bolsonaro seems to represent a dynamic of legitimization that does not end with the individual, but with the apparently effective and profuse activity among those who reproduce his discourses.

Moreover, the presence of pro-vaccine movements in the conversation is undeniable. In fact, the science communicator Átila Lamarino and the public health doctor Daniel Dourado were among those the most relevant in the sample analyzed. Although users who shared their content were the least active on the platform, they were more diverse, and from what we could analyze, also more influential in their network.

Intercom, Rev. Bras. Ciênc. Comun., São Paulo, v. 46, e2023121, 2023

17/21

In terms of the topics covered, the conversation about vaccines and vaccination policies on Twitter was wide-ranging. Accountability and Public Policies were the most prominent, partly as a result of the participation of the most influential profiles (Atila, Accountability, Jair Bolsonaro, Public Policies).

Related to this, the presence of themes such as Bioethics, Personal Experiences and Economic Aspects show that Twitter presents itself as a privileged locus for health studies, because of the possibility of analyzing perceptions, experiences and opinions regarding health policies, programs and services.

Finally, it is important to mention that we could not deepen the analysis on all aspects related to disinformation in the conversation, especially on topics such as Vaccine Safety and Effectiveness, Conflicts and Conspiracy Theories, so there is a need for further work addressing this problem, given its importancefor Public Health.

References

1 em cada 4 brasileiros pode não se vacinar contra a COVID-19. AVAAZ, 7 set.2020. Available at: < https://secure.avaaz.org/campaign/po/brasileiros_nao_vacinar_covid/>. Access on: 6 abr. 2021

ABD-ALRAZAQ, A. et al. Top Concerns of Tweeters During the COVID-19 Pandemic: Infoveillance Study. **Journal of medical Internet research**, v. 22, n. 4, 1 abr. 2020.

ABUZAYED, A.; AL-KHALIFA, H. BERT for Arabic Topic Modeling: An Experimental Study on BERTopic Technique. **Procedia Computer Science**, v. 189, p. 191–194, 1 jan. 2021.

ALAGHA, I. Topic Modeling and Sentiment Analysis of Twitter Discussions on COVID-19 from Spatial and Temporal Perspectives. **Journal of Information Science Theory and Practice**, v. 9, n. 1, p. 35–53, 1 mar. 2021.

AMARAL, I.; SANTOS, S. J. Algoritmos e redes sociais: a propagação de fake news na era da pós-verdade. **As fake news e a nova ordem (des)informativa na era da pós-verdade**, p. 63–85, 1 jan. 2019.

BOON-ITT, S.; SKUNKAN, Y. Public perception of the COVID-19 pandemic on twitter: Sentiment analysis and topic modeling study. **JMIR Public Health and Surveillance**, v. 6, n. 4, p. e21978, 1 out. 2020.

CINELLI, M. et al. The COVID-19 social media infodemic. **Scientific Reports**, v. 10, n. 1, 1 dez. 2020.

DAHLGREN, P.; ALVARES, C. Political Participation in an age of Mediatisation. **https://doi.org/10.1080/131 83222.2013.11009114**, v. 20, n. 2, p. 47–65, 2014.

DEUZE, M. What is journalism? Professional identity and ideology for journalists reconsidered. Journalism, 6(4), 442-464, 2005.

EYSENBACH, G. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. **Journal of medical Internet research**, v. 11, n. 1, 1 jan. 2009.

FIGUEIREDO, A. et al. Mapping global trends in vaccine confidence and investigating barriers to vaccine uptake: a large-scale retrospective temporal modelling study. **The Lancet**, v. 396, n. 10255, p. 898–908, 26 set. 2020.

GROOTENDORST, M. BERTopic: Neural topic modeling with a class-based TF-IDF procedure. 11 mar. 2022.

HAMAN, M. The use of Twitter by state leaders and its impact on the public during the COVID-19 pandemic. **Heliyon**, v. 6, n. 11, p. e05540, 1 nov. 2020.

KANG, Gloria J. et al. Semantic network analysis of vaccine sentiment in online social media. **Vaccine**, v. 35, n. 29, p. 3621-3638, 2017.

LAVAZZA, A.; FARINA, M. The Role of Experts in the Covid-19 Pandemic and the Limits of Their Epistemic Authority in Democracy. **Frontiers in Public Health**, v. 8, p. 552192, 14 jul. 2020.

LIMA, C. R. M. DE et al. Emergência de saúde pública global por pandemia de Covid-19. **Folha de Rosto**, v. 6, n. 2, p. 5–21, 11 maio 2020.

MELO, T.; FIGUEIREDO, C. M. S. Comparing news articles and tweets about COVID-19 in Brazil: Sentiment analysis and topic modeling approach. **JMIR Public Health and Surveillance**, v. 7, n. 2, p. e24585, 1 fev. 2021.

MONARI, A. C. P. et al. Legitimando um populismo anticiência: análise dos argumentos de Bolsonaro sobre a vacinação contra Covid-19 no Twitter. **Instituto Brasileiro de Informação em Ciência e Tecnologia**, v. 17, n. 1, p. e5707–e5707, 2 jun. 2021.

MONARI, P.; BERTOLLI FILHO, C. Saúde sem Fake News: estudo e caracterização das informações falsas divulgadas no Canal de Informação e Checagem de Fake News do Ministério da Saúde. **Mídia e Cotidiano**, v. 13, n. 1, p. 160–186, 26 abr. 2019.

OLIVEIRA, T. et al. Politização de controvérsias científicas pela mídia brasileira em tempos de pandemia: a circulação de preprints sobre Covid-19 e seus reflexos. **Revista Brasileira de História da Mídia**, v. 10, n. 1, p. 30–52, 20 jul. 2021.

PASCHOALOTTO, M. A. C. et al. Running away from the jab: factors associated with COVID-19 vaccine hesitancy in Brazil. **Revista de Saúde Pública**, v. 55, p. 97–97, 26 nov. 2021.

RECUERO, R. et al. **Desinformação, Mídia Social e Covid-19 no Brasil: Relatório, resultados e estratégias de combate**. Pelotas: [s.n.].

RECUERO, R.; SOARES, F. B. O Discurso Desinformativo sobre a Cura do COVID-19 no Twitter: Estudo de caso. **E-Compós**, v. 24, p. 1–29, 27 jul. 2021.

RECUERO, R.; ZAGO, G. A Economia do Retweet: Redes, Difusão de Informações e Capital Social no Twitter. **Revista Contracampo**, n. 24, p. 19–43, 7 ago. 2012.

SILVEIRA, R. et al. Topic Modelling of Legal Documents via LEGAL-BERT 1. 2021.

SINGER, J. B. User-generated visibility: Secondary gatekeeping in a shared media space. New media & society, v. 16, n. 1, p. 55-73, 2013.

Vacina contra covid-19: Reino Unido é 1º país a aprovar imunizante da Pfizer e pode iniciar aplicação em uma semana. BBC News Brasil, 2020. Available at: < https://www.bbc.com/portuguese/internacional-55156721>. Access on: 10 abr 2021.

VIS, F. A critical reflection on Big Data: Considering APIs, researchers and tools as data makers. **First Monday**, 2 out. 2013.

WORLD HEATH ORGANIZATION. **Ten threats to global health in 2019**. Available at: <https://www.who.int/news-room/spotlight/ten-threats-to-global-health-in-2019>. Access on: 23 nov. 2022.

YIN, H. et al. Sentiment analysis and topic modeling for COVID-19 vaccine discussions. **World Wide Web**, v. 25, n. 3, p. 1067–1083, 1 maio 2022.

Intercom, Rev. Bras. Ciênc. Comun.,
São Paulo, v. 46, e2023121, 2023

19/21

YOUSEFINAGHANI, S. et al. An analysis of COVID-19 vaccine sentiments and opinions on Twitter. **International Journal of Infectious Diseases**, v. 108, p. 256–262, 1 jul. 2021.

ZAGO, S. O Twitter como Suporte para Produção e Difusão de Conteúdos Jornalísticos. **C-Legenda - Revista do Programa de Pós-graduação em Cinema e Audiovisual**, n. 21, 17 dez. 2009.

About the authors

Arthur da Silva Lopes
Interdisciplinary Bachelor in Health from the Federal University of Bahia (UFBA) and Data Scientist. She is a master's student in the Postgraduate Program in Collective Health (ISC/UFBA) and a researcher in the Media and Health axis of the Observatory of Political Analysis in Health (ISC-UFBA). She is also a collaborator at the National Institute for the Public Communication of Science and Technology (INCT-CPCT), Oswaldo Cruz Foundation (Fiocruz). E-mail: arthur.lopes@ufba.br

Antonio Marcos Pereira Brotas
He has a PhD in Culture and Society from the Federal University of Bahia (UFBA), a master's degree in Contemporary Communication and Culture from UFBA and a post-doctorate from the Post-Graduate Program in Science, Technology and Health Communication at Casa Oswaldo Cruz (COC/Fiocruz). He is currently a public health technologist at the Gonçalo Moniz Institute (IGM - Fiocruz Bahia), journalist, communications advisor, permanent lecturer on the Master's Degree in Science, Technology and Health Communication (COC/Fiocruz) and researcher at the Global Health and Neglected Diseases Research Laboratory (LISD). E-mail: brotas@bahia.fiocruz.br

Luisa Massarani
She is the coordinator of the National Institute for the Public Communication of Science and Technology and a researcher at the Casa de Oswaldo Cruz, Oswaldo Cruz Foundation. She teaches in the Postgraduate Program in Teaching in Biosciences and Health at the Oswaldo Cruz Institute/Fiocruz, in the Postgraduate Program in Teaching at the Institute of Medical Biochemistry at the Federal University of Rio de Janeiro and in the Academic Master's Program in Dissemination of Science, Technology and Health at the Casa de Oswaldo Cruz. She has a degree in Social Communication from the Pontifical Catholic University of Rio de Janeiro (1987), a master's degree in Information Science from the Brazilian Institute of Information in Science and Technology (1998), a PhD in Management, Education and Dissemination in Biosciences from the Federal University of Rio de Janeiro (2001), a post-doctorate at University College London (2013) and a post-doctorate at Oregon State University (2015-2016). She has published around 230 articles in national and international scientific journals. E-mail: luisa.massarani@fiocruz.br

20/21

Intercom, Rev. Bras. Ciênc. Comun.,
São Paulo, v. 46, e2023121, 2023

Authors' contribution

Lopes, A. S: research, conceptualization, methodology, data curation, formal analysis, writing - original draft, writing - proofreading and editing, and approval of the final version of the manuscript; Brotas, A. M. P.: acquisition of funding, conceptualization, supervision, writing - proofreading and editing, and approval of the final version of the manuscript. Massarani. L.: acquisition of funding, writing - proofreading and editing, and approval of the final version of the manuscript.

Availability of data

The authors confirm that data supporting the research is contained in the article and/or supplementary material.

Conflict of interest

There is no conflict of interest.

Editorial data

Intercom, Rev. Bras. Ciênc. Comun.,
São Paulo, v. 46, e2023121, 2023

21/21