Characterizing The Impact Of Fact-Checking On The COVID-19 Misinformation Combat

Cefas Garcia Pereira
Department of Computer Science
Pontifical Catholic University of Minas Gerais
(PUC Minas)
Poços de Caldas, Minas Gerais
cefas.pereira@sga.pucminas.br

Humberto Torres Marques-Neto
Department of Computer Science
Pontifical Catholic University of Minas Gerais
(PUC Minas)
Belo Horizonte, Minas Gerais
humberto@pucminas.br

ABSTRACT

The COVID-19, a disease caused by SARS-CoV-2, affected the whole world in 2020 by its pandemic impact. This virus has a very high capacity for contamination through contact with other infected people. One of the main ways to fight the virus is to reduce the possibility of contact with the infected population by avoiding the crowding of people. Within this context, the virtual means of communication are being channels of information about the pandemic and also the externalization of users' feelings and opinions. Through social networks, people assume the role of content generators and not just consumers. This leaves room for the spread of misinformation, biased news, and rumors that are originated from laymanship, political and commercial interests. This work aims to characterize how fact-checking agencies have reacted in the combat against false information about COVID-19 on social networks such as Twitter and Facebook, seeking to broaden the understanding of misinformation propagated over the internet. During the study, we collected fact-checking articles about COVID-19 written by experts from different countries. Through the verified news, we searched social media posts which misinformation began to be spread. After collecting this data, it was verified how long it took the fact-checking agencies to analyze the veracity of the news. In addition, the texts were processed to detect whether the topics being dealt with by the agencies are, in fact, those with the greatest engagement of users within the analyzed social networks, and also the presence of bots on social media. We compared the collection of fact-checking provided by the Poynter Institute and Google's Fact-Checking API, to identify a uniformity between the databases. The results showed that the response time of agencies was around 23 days in the case of misinformation on Twitter and approximately 6 days on Facebook.

CCS CONCEPTS

• **Human-centered computing** → **Collaborative and social computing devices**; • **Information systems** → **Document topic models**;

SAC '22, April 25–29, 2022, Virtual Event,

ACM ISBN 978-1-4503-8713-2/22/04.
https://doi.org/10.1145/3477314.3507183

KEYWORDS

Fact Checking; Misinformation; Coronavirus, COVID-19; SARS-CoV-2; Social Network, Twitter; Facebook.

ACM Reference Format:
Cefas Garcia Pereira and Humberto Torres Marques-Neto. 2022. Characterizing The Impact Of Fact-Checking On The COVID-19 Misinformation Combat. In *The 37th ACM/SIGAPP Symposium on Applied Computing (SAC '22), April 25–29, 2022, Virtual Event,* . ACM, New York, NY, USA, Article 4, 8 pages. https://doi.org/10.1145/3477314.3507183

1 INTRODUCTION

The year 2020 was marked by the COVID-19 (Coronavirus Disease 2019) pandemic, which, in 2021, is still impacting the whole world. The SARS-CoV-2 virus, also called Coronavirus, has been classified as a serious global threat by the World Health Organization [20]. The official statement dates the virus to appear in December 2019, and it was declared a pandemic in March 2020. After the dramatic increase in the number of occurrences of respiratory and lung problems in the city of Wuhan, China, researchers isolated the new virus responsible for the contamination during the patients' treatment. [7]. Early on, countries like Italy and Spain were heavily affected by the pandemic, with record rates of contamination and death from COVID-19. The United States, Iran, Germany, and Asian countries also quickly suffered from the effects of Coronavirus.

After the WHO declaration in March 2020, countries started to adopt measures to prevent the virus. Many of them began to adopt social isolation, blocking their borders and restricting the movement of people. Businesses have largely closed, keeping only supermarkets, pharmacies, bakeries, and other essential services open. The world had to adapt to this new routine and as a result of social isolation, people started to carry out their activities such as education and work online, usually by their homes.

The information and social relationship through the virtual space, which has naturally been growing over the years, has become almost an exclusive possibility. People started to get informed and also to express their opinions, even more through the internet, especially on online social networks such as Facebook, Twitter, Instagram, and WhatsApp. In fact, not only social networks but the internet as a whole has enormous power to contribute in situations like this, of social isolation resulting from a pandemic. The speed in the dissemination of information, the simplicity in which it is transmitted, and also the countless number of options for textual form and style make people find how they identify themselves more and so they become much more engaged and informed. Several researchers showed that social networks play an important role as a source of information for people and that they also help to understand public behaviors during important events [1, 8].

However, the presence of technology in people's lives brings positive and negative effects, because, at the same time it democratizes access to information, it creates an opportunity for political and commercial manipulation and the dissemination of fake news. The fight is not exclusive to the virus, but also people's misinformation. Conspiracy theories can not only nullify the effect of the serious work being done to fight the coronavirus, but also make it worse.

This work aims to understand the way fact-checking agencies have behaved in this scenario, especially through the identification of the time when the agency verifies the information and how this can impact the engagement of users. As well as analyzing the fake content that is shared on social networks about the COVID-19 pandemic and its interaction by users. Understanding these issues is important, as misinformation has the power to worsen the effects of events such as the outbreak of COVID-19, therefore, it is necessary to understand how combat is currently carried out to further mitigate the effects of misinformation. Therefore, this work strives to characterize the reaction of fact-checking agencies to posts about COVID-19 on Facebook and Twitter.

This study, therefore, contributes to the proposal of a methodology to understand the impact of fact-checking on the fake news that circulates in social networks for a given subject. The work also contributes with the results generated by the analyses, demonstrating that fact-checking agencies take approximately 6 days to verify information on Facebook and around 23 days on Twitter. Geographically, India, the United States, Brazil, and Spain were detected as hot spots concerning the spread of misinformation about the pandemic. The research also showed that topics such as cure, vaccine, and treatment are common in the two social networks analyzed and that, fortunately, they have been fought by fact-checking agencies, as these same issues were identified in the articles written by the agencies. It was also identified that most of the posts were made by human authors, that is, relatively few *bots* were identified. Unlike previous studies that analyzed public discourse during the pandemic, this work focused on discoveries directed at the dissemination of false information and the activities of fact-checking agencies. A comparison was also made between Poynter's fact-checking basis and the Google Fact-Checking API to assess the consistency between them in terms of the number and frequency of posts.

This work is organized as follows. Section 2 presents the previous works related to misinformation, approaching the main concepts on which this research is based. Section 3 discusses the methodology, which includes data collection and pre-processing used for the analyses. In Section 4 the analyzes and results are shown and debated. Limitations are pointed out in Section 5; the conclusion is presented in Section 6.

2 RELATED WORK

We consider here that the term "misinformation" is designated as information that does not accurately reflect the true state of the news [10]. The term "misinformation" is applied to information that is initially presented as true and which is later shown to be false. The causes and effects of misinformation are still study subject of several researchers. Although misinformation is not a recent phenomenon [17], the digital media, due to its ability to quickly transmit a large amount of data, has brought this issue to the fore.

The reasons for the spread of misinformation are diverse. Research [9] claims that a category of misinformation originates from the scientific community, namely "bad science". The author exemplifies her position through the 1998 publication, by the then physician Andrew Wakefield, which intended to show a link between autism and the measles, mumps, and rubella vaccine. Despite having his license revoked and his work withdrawn, Wakefield insists on his anti-vaccination campaigns. Larson argues that if Wakefield had been disciplined and his article withdrawn 12 months after publication, rather than 12 years, many health-related problems and distrust of the vaccine could have been avoided.

Work carried out [5] revealed the use of *bots* (robots) to disseminate hate speech associated with anti-vaccination movements aimed at segregating political groups in Russia, which leads to another category of misinformation, the category moved by political interests. There are also cases in which misinformation is the result of news that has certain veracity in the facts, but is distorted due to exaggeration, coincident sequence of facts, misunderstanding, or common sense [9].

Work carried out [5] revealed the use of *bots* (robots) to disseminate hate speech associated with anti-vaccination movements aimed at segregating political groups in Russia, which leads to another category of misinformation, the category moved by political interests. There are also cases in which misinformation is the result of news that has certain veracity in the facts, but is distorted due to exaggeration, coincident sequence of facts, misunderstanding, or common sense [9].

It is essential, therefore, that ways to identify and combat misinformation are developed. It is difficult for a permanent solution to exist, but some measures can be taken to curb the growth of this phenomenon. Some social networks have already started to adopt moderation measures based on complaints or content verification by experts.

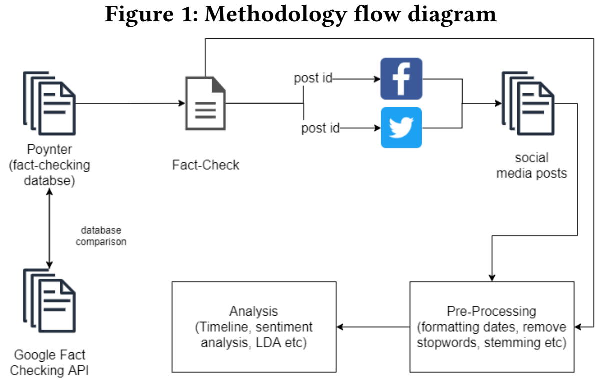
One of the tools used to combat false information is fact-checking. Its great mission is to raise the cost of lying, that is, not let the process of spreading misinformation be so trivial. There is still great difficulty in performing the automatic identification of misinformation or fake news. Many studies are being carried out to solve this challenge, but so far, there is no consensus on the best way to detect misinformation. Fact-checking is generally carried out by journalists who seek, through new sources, research, interviews, and other journalism resources, to verify the veracity of information [18]. The methodology adopted by the agency is an individual decision, in which journalists prioritize the issues according to the criteria they deem most relevant, as well as how the information circulates and its repercussion. Fact-checking agencies does not take into account opinions, predictions and does not point out trends. The agency must reserve its efforts to validate the veracity of the facts. Upon completion of the fact-checking, the journalist writes a story explaining the facts and justifying the classification (false, true, exaggerated, etc.) of the verified information [18].

The role of fact-checking has gained increasing importance since in the context of digital communication, evaluating the quality of public discourse can raise the quality of information circulating in networks and provide society with more accurate data. This work chose to use the *Poynter* database as a starting point to identify misinformation, since the aim of the work is not to propose an automatic system for predicting misinformation, but rather to understand the interaction of users with misinformation on social networks and how fact-checking can impact them. Unlike previous studies that analyzed public discourse during the pandemic, this work focused on discoveries directed at the dissemination of false information and the activities of fact-checking agencies.

3 METHODOLOGY

To achieve the goals, we performed three steps as described below. The first step is related to data collection. This step includes identifying the topics that are relevant to search on social media and also the way in which data will be collected. We started identifying the topics, through the fact-checking database provided by Poynter, which is an independent journalism agency specialized in fact-checking, which offers greater security to research, since it is possible to assume that the misinformation collected has already been verified by a team that masters this technique. Subsequently, work continues on preparing the data. This means that the data is treated in such a way that it meets the restrictions of the tools and algorithms that will be used for the analyses. Finally, with the data collected and processed, the next step is to analyze them. The analyzes consist of metrics that reveal the topic's popularity, emotional impacts, and user engagement. Figure 1 illustrates the phases from identifying the topics that originate the search terms, to analyze the posts.

Figure 1: Methodology flow diagram



3.1 Fact-Checking

The *Poynter* institute, founded by Nelson Poynter in 1975, has played an important role during the pandemic period. In 2015, the institution opened its IFCN unit to promote and make available fact-checking in the world. In 2020, the unit made available an online database with fact-checking on Coronavirus from several countries [13]. Through a web interface, any user can explore the database and verify information about the pandemic. Access to this type of information is extremely important as it empowers the reader to check facts quickly and easily.

This database works as a centralizer of the fact-checking that has been carried out around the world, it was used in this work to collect the materials written by the agencies. A *crawler* was developed using the programming language *Python* [1]. Crawlers are programs that simulate human access to web pages. Initially, the entire universe of accessible articles was collected. The collection in the base of *Poynter* returned 11,666 publications from fact-checking. The fact-checks were carried out by 99 different agencies and the verified information circulated to 130 countries; 37 labels were identified, including true, false, misleading, unproven, imprecise, etc.

3.2 Social Networks: Facebook and Twitter

After the first collection, posts were collected on social networks. For this, the strategy adopted was to check the body of the fact-checking article for the presence of posts on the social networks Facebook and Twitter, chosen because of its high popularity in different countries, especially in the content of *Poynter*. The crawler went through all the hyperlinks of the fact-check and filtered out those from Facebook and Twitter, so a fact-check can have more than one post associated with it.

The hyperlinks from social networks can come in different formats, so scripts with strings manipulation and regular expressions were executed to extract the *IDs* from the posts. Also, it was identified that some agencies host the links of social networks posts on *archive.today*, which is a kind of server that guarantees that the content will remain online even if it is deleted in the original source. The collection also considered the links contained in the *archive.today* and retrieved them to ensure as much material as possible.

Subsequently, the data was filtered. The article labels were grouped so that only false information were analyzed, as the focus here is on misinformation. Therefore, the labels placed in front have all been converted to "*False*": *False, False and Without Context, false context, false headline, False, MISLEADING/FALSE, Incorrect, Mainly False, Mostly False , NOT LEGIT, misleading*.

After this grouping, the articles were filtered by the label "*False*", leaving 10,170 (87.17%) of the initial collection. It was possible to identify how many fake articles from *Poynter* contain the *links* for social networks. Of the 10,170 filtered articles, it was identified that 42% refer to some social network post analyzed in this work (Twitter 21%, Facebook 16% and 5% in both). Thus, we can say that some agencies do not insert the redirect link for the post on Facebook and/or Twitter or that 58% of the news did not actually circulate on these social networks, being broadcast exclusively on other closed social networks such as WhatsApp, on TV, newspaper, magazines, radio, blogs or even other media.

We also identified that some agencies only posted images (*screenshots*) of the post. In an attempt to leverage this content, we used an OCR – *Optical Character Recognition* to try to recognize text in images. However, the results were not promising. The code written in *Python*, made use of the *open-source Tesseract, Microsoft Computer Vision API* and *Google Vision* libraries. None of the alternatives were able to accurately extract the image content due to the low quality

[1] https://github.com/cefasgarciapereira/fact-checking

and lack of standard of the images. Thus, this approach must be explored in future work.

We, therefore, decided to proceed with the analysis of the 42% of news collected by crawler on *Poynter*. The dataset was added to the *IDs* of the posts from social networks and submitted to another collection, responsible for extracting the metadata of the posts. In this phase, the open-source *Tweepy* library was used, which is an abstraction of the official Twitter API for Python. For Facebook, the collection was implemented by the authors themselves, also through crawlers. The collection resulted in a total of 2,908 posts on Facebook and 4,096 on Twitter. It is important to note that these posts are those that are still online, as those that were banned or deleted can no longer be captured. It was even possible to detect that 931 posts (32%) on Facebook are marked with a false information alert, even though the content is available. This measure already shows that there is a fight on the part of the social network against this type of content. On Twitter, this analysis was not possible, as the library does not return this information.

3.3 Pre-Processing

At the time of collection, some techniques had already been applied so that the data could be extracted and structured. However, for the analysis to be possible, it was necessary to perform treatments on the collected data. Starting with the dates, since each platform uses different kinds of representation, the data were submitted to pre-processing written by the authors, to standardize the dates. Text content has also undergone standardization. Those that do not belong to the English language were translated programmatically by the *Google Translate* API. Afterward, the cleaning step was performed, in which special characters and *stopwords* were removed. Also, the texts were divided into words, a process called *tokenization* and finally went through the step of *stemming*, which is the process of grouping the inflected forms of a word so that they can be analyzed as a single word.

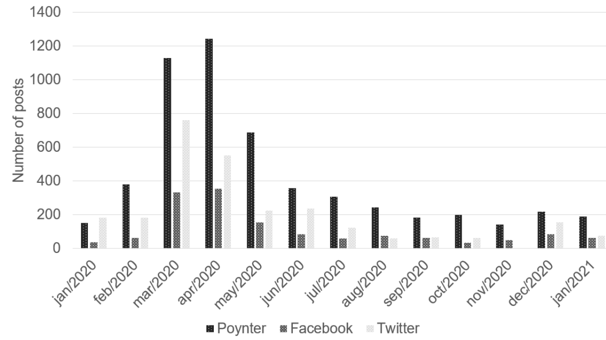


Figure 2: Number of posts per month on each platform

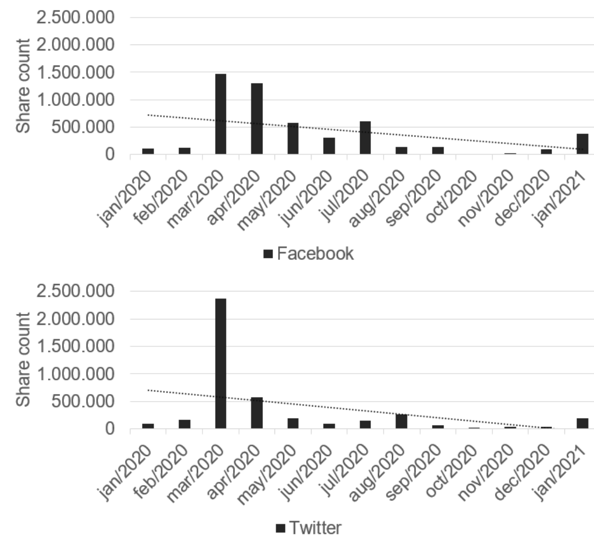


Figure 3: Average share per month on each social network

4 ANALYSIS AND RESULTS

With the analysis of the generated datasets, it was possible to see the behavior of publications throughout the pandemic period. Publications range from January 2020 to January 2021, as shown in Figure 2. We can see that the curve over the months is extremely similar for all platforms. Note that the period with the highest number of publications was between March and May 2020, that is, at the beginning of the COVID-19 pandemic. A positive indication regarding the reaction of fact-checking agencies is that the number of articles was also more intense at the beginning of the pandemic.

Figure 3 shows that the average share, as well as the number of publications in Figure 2, was higher in the first months of the pandemic but follows a downward movement, as shows the linear trendline of the graph. It is important to note that on Facebook, the sharing average for the month of January 2021 was the highest since July 2020, which may indicate an increase in the number of publications for the first half of 2021.

To find out how long it takes for an agency to identify the information on Facebook and/or Twitter, verify it, and publish the fact-checking, the difference between the publication date of the fact-checking and the date of publication of posts on Facebook and/or Twitter was calculated.

The results, presented in Figure 4, reveal that, on average, agencies took 23 (SD=33.6) days to publish the verification of the information shared on Twitter and 6 (SD=7.04) days to Facebook. The verification time was calculated through the difference between the day of publication of the fact-checking and the day the post was published on the social network. The high value of the standard deviation and a large number of outliers allows us separate the cases in which the agency's response to the post on the social network was quick and the cases in which the response took a long time to occur. A value treatment was applied to remove outliers. The technique chosen was the Tukey method, as it was recommended in past works [3]. However, to understand the phenomenon of outliers, we manually inspected a portion of the posts and we identified that some subjects were discussed for a long time, and therefore, the fact-checking matters revisited some posts. For example, as the vaccine is a recurrent issue, posts on the subject "vaccine" at the beginning of the pandemic were referenced in fact-checking from 2021 and this justifies the long term in these cases.

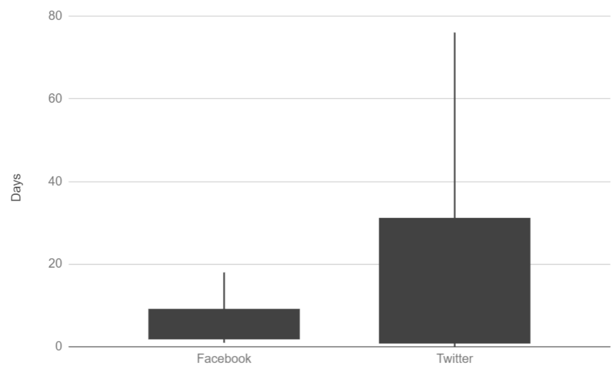


Figure 4: Difference in the number of days between the date of posting on the social network and the date of publication of the fact-checking without outliers

Another situation that occurs is when a social network user posts old content, however, manipulating some information to bring it into the context of the pandemic. For example, a video post that has been shared several times on Facebook and Twitter claims to show a mock funeral procession to "celebrate" Indian Interior Minister Amit Shah, who tested positive for the coronavirus. The claim, however, is false as the video actually shows a simulated funeral procession during a protest against the anti-citizenship law in the Indian state of West Bengal in 2019 [2]. For this reason, the 2019 post appeared in an August 2020 fact-check and therefore resulted in a large difference of days between the fact-checking and the social media post.

As for the geographical disposition, we took into account how much information was verified as false by the fact-checking agencies in each country. Figure 5 shows that India, Brazil, the United States, and Spain are hot areas in the number of circulation of false information. It is important to point out that India and the United States are countries that natively use english to communicate and this can easen the verification of information by agencies around the world that may not find the same ease in other languages, and as a consequence, these two countries are at the top of the list, as shown in Figure 5.

Countries	Number of fact checks
USA	1040
Brazil	460
France	329
India	300
Turkey	222
Germany	212
Colombia	194
Taiwan	124
Spain	114
Philippines	102

Figure 5: Hot spots in misinformation dissemination

Anyway, the agencies are usually careful to check the circulation of information in other countries, as is the case with the information verified by AFP, an alleged video in which a woman trembling while walking, it would be due to the side effects of the vaccine from *Pfizer* [11]. However, no evidence was found by experts that could link the two facts. It was found that this rumor circulated in the United States, Brazil, and India; therefore, language is not a limiter, but indirectly it may favor a larger number of cases in those English-speaking countries.

To extract greater knowledge about the topics discussed on social networks and the information verified by fact-checking agencies, the pre-processed texts were submitted to analysis using a term frequency (unigram) and the pairs-term frequency (bigrams). To visualize this analysis, we generated word clouds as shown in Figure 6. The word cloud can give us visibility to the words that most appeared in the set of documents and this helps to identify the most discussed subjects.

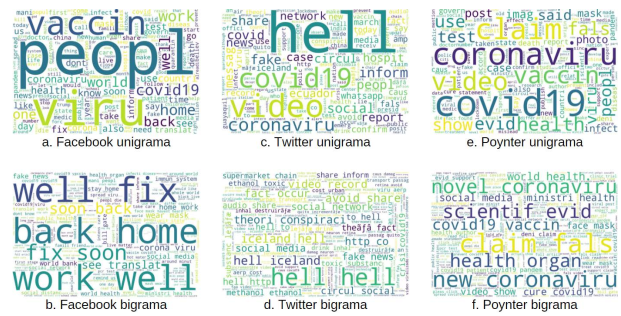


Figure 6: Word frequency and word pair clouds per document set

To identify the topics, LDA (*Latent Dirichlet Allocation*) was applied [4], a widely used topic modeling algorithm [1, 6]. Topic modeling is an unsupervised machine learning technique that can locate groups in a collection of text documents. The *Gensim* library for *Python* was used to implement the code responsible for processing the texts. The LDA receives as a parameter a fixed set of topics. Each topic is represented by a set of words. The purpose of LDA is to map the documents provided to the set of topics so that the words in each document are mainly associated with those topics defined by the algorithm. This work used modeling to find topic groups in text content from Facebook, Twitter, and *Poynter* posts. For the modeling, it is necessary to generate a *corpus*, which is a representation of the words and their frequency of occurrence in the set of documents. One *corpus* was generated for Facebook and another for Twitter, in both cases the text contained in the publication was used for this purpose. In *Poynter*, the *corpus* is the agency's explanation text, which is found in the "*justify*" field of the collected dataset.

Before applying the modeling, tests were performed to assess the level of coherence between the words and their respective topic. This test was applied to determine which number of topics is the most coherent for each set of documents. The topic coherence measure scores a single topic by measuring the degree of semantic similarity between the highest valued words within the topic. A set of statements or facts is considered coherent if they support each other. Thus, a coherent set of facts can be interpreted in a context where it covers all or most facts [14]. We used the Gensim's implemented version of *u_mass* that returns positive and negative values in a way that the closer to zero the better [15].

After testing, LDA was run with 6 topics for Twitter, 7 for Facebook, and 6 for Poynter. These numbers determined as parameters were chosen based on the coherence test and do not lead us to an arbitrary decision. Subsequently, we manually associated the topics retrieved by LDA with the word clouds and reached a consensus on the set of relevant topics for each collection of documents. Then, a rule-based classification script written in Python was used to verify the presence of any words that make up a pre-identified topic in each collection of documents. The script is a simple matching technique to see if a given text contains keywords extracted from one of the LDA topics. A text that contains a keyword related to a certain topic was classified as belonging to that topic and, therefore, a text can be evaluated as belonging to more than one simultaneous topic.

To find out the level of user engagement for each topic, the total number of interactions with the post (likes, comments, and shares) was added up and divided by the total number of posts belonging to the same topic. In the case of *Poynter* articles, as there is no user interaction, only the number of occurrences was considered, that is, how many articles belong to each topic. Thus, we arrived at the values contained in the Table 1, which show the topics identified for each context and the level of interaction for each topic, which we called "engagement level". It is important to note that the values contained in the engagement level column are not comparable among different datasets, that is, comparing the engagement level of a Twitter topic with Facebook would not make sense, as these are another kind of interactions, universe of users, topics, etc. This value is important to show which subjects users reacted the most within each social network and in the case of *Poynter*, which topic had the highest number of publications by agencies.

Documents Set	Topic	Words	Engagement Level
Twitter	Fake News	conspiraci, fake news, false, suppos	1,989,888
	Russian vaccine	russia, vaccin, president, putin	1,036,230
	India	vandana, gehlot, ashok, chouhan shivraj, delhi	1,723
	Humanitarian Crisis	crise, migrant, hospital, water, crisi, infect, peopl, health, protest	2,251,289
Facebook	Vaccine, Cure and Treatment	vaccin, mask, test, viamin	7,517.96
	Politics	trump, govern, china, economy, politic, presid	5,269.31
	China	china, wuhan	31,539.90
	Quarantine	lockdown, outbreak, home	1,804.30
	Deaths	death, died	378.46
Poynter	Fake News	false, mislead, fake, claim	3,117
	Vaccine, Cure e Treatment	vaccin, cure, mask, vitamin	1,981
	Bill & Melinda Gates	bill, melinda, gates	370
	Deaths	death, dead, die	1,148
	Quarantine	outbreak, lockdown	471

Table 1: Identified topics, by set of documents, and their level of engagement

It is possible to notice that discussions around methods to reduce the disease, vaccine, cure, and treatment appeared with a lot of evidence in the document collections, as well as issues related to the *fake news* itself and the dissemination of misinformation. The topic involving deaths also proved to be relevant. In the case of Twitter, the topic India emerged as a result of a set of names of Indian personalities and also the name of the capital Delhi. Discussions involving politics also surfaced, in the case of Facebook, through relations between the United States, China, and former president Donald Trump; on Twitter this issue appeared in the form of humanitarian crisis, water crisis, protests, problems involving immigration, and the situation of hospitals. In *Poynter*, the association of terms around the last name Gates brought to light rumors involving the Gates foundation of Bill and Melinda Gates. Several rumors involving the Gates foundation with government funding, vaccine development, and commercial deals involving the foundation have been identified by fact-checking agencies. China was also one of the emerging topics on social media, especially on Facebook, which is strongly associated with the origin of the virus and the pandemic.

Seven main topics that are being discussed recurrently on social networks were identified. They are the very dissemination of *fake news* and conspiracy theories, Russian vaccine, politics, quarantine, deaths, origin of the virus, and finally the cure and treatment of the disease. The word clouds revealed a strong sense of panic, it is possible to detect terms and posts that talk about quarantine, locks, and requests to stay at home, on Twitter the word "*hell*" appeared prominently and it reflects this feeling well. Many users are reluctant to believe in the fact of the pandemic and it was possible to identify a sensationalist and distorted use of social networks. Others seem to be looking for alternative solutions to the pandemic, unscientific explanations of supposed cures or ways to fight the virus. Still, conspiracy theories disseminated through decontextualized videos and images covertly defend the effects of the vaccine or that the virus does not exist.

This is a great challenge for public health, which must make efforts to fight not only the pandemic but also the "infodemic" [6]. The dissemination of false information on social networks can fuel widespread panic and cause harmful effects to the population, blurring the evidence and making it difficult for professionals and public health systems to respond [1, 19].

In this work, the profiles of the authors of posts verified by fact-checking agencies were submitted to the Botometer platform to identify the possible presence of *bots* in the publications collected on Twitter [16]. Based on the aforementioned criteria, the algorithm returns a generalized score from 0 to 1, so that the lower the value, the less chance that the profile is a *bot* and the higher the value, the greater the chance that the profile is a *bot*. The results suggest that the presence of *bots* was low in the set of publications analyzed by the agencies, as 75% of users received a score at the bottom of the set, that is, below 0.5. Only 11% of profiles have a score above 0.8, which indicates a high probability that the profile is a *bot*. The results of the evaluation carried out that the *bots* are more used to promote and disseminate the content than as a generator role. As the checks carried out by *fact-checking* agencies are focused on the source of the misinformation, most of the time, the profiles humans.

To extend and broaden the understanding of the set of fact-checks materials available on the internet, a comparative study was carried out between two fact-checking database. The Poynter database, used in the previous analysis, and the Google database, available through an API[2] and in a public file format[3].

The first step was to understand the universe of fact-checking articles about the COVID-19 pandemic. Google’s *dataset*, as shown in Figure 7, has 28,900 *fact-checking* articles on various subjects, so it was necessary to filter out those linked to the COVID-19 pandemic . The filter was done using words contained in the titles and corresponding URLs of each fact-check. The set of words used in the filter was inspired by previous works [1, 6] and contains the following terms: covid, corona, sars, coronavirus outbreak, cov2, ncov, Wuhan virus, and china virus.

As Figure 7 illustrates, from the total Google set, about 3,561(12%) stories correspond to the pandemic. Of these, 3,161 fact-checking stories are contained on both platforms. Poynter, with its database dedicated exclusively to the *fact-checking* of the coronavirus, has significantly more articles available on its platform, with 11,666 articles that verified facts about the coronavirus pandemic.

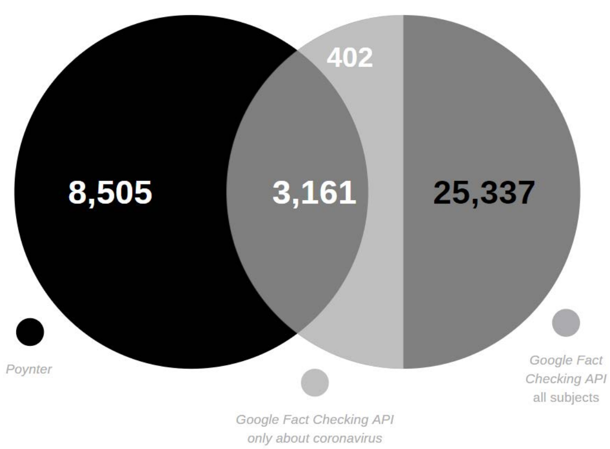


Figure 7: Venn diagram of fact-checks contained in Poynter and Google Fact-Check Tool API

As for the timing of posts, they have very similar distributions. The same pattern noticed earlier is repeated for the fact-checking available on Google, that is, a sharp movement in the first half of 2020, followed by a sequence of falls in the following months. Another analysis carried out was to verify which fact-checking agencies had greater importance in the number of publications that appeared in each database. It was found that the AFP agency received a very high number of fact-checks, with a total of 1919 publications, considering the two platforms, it represents about 16% of the volume of fact-checks. AFP was launched in 2017 in France as a digital fact-checking service and eventually expanded to five continents. In second place in the ranking is Politifact, another extremely important agency in this medium, which began its activities in 2007 as an election-year project by the Tampa Bay Times, a daily newspaper in Florida. In third place is Maldita.es, a Spanish institution that has several branches dedicated to monitoring public discourse [12].

[2]https://toolbox.google.com/factcheck/explorer
[3]https:// www.datacommons.org/factcheck/

5 LIMITATIONS

To achieve the objective of the work, it was necessary to explore information that was confirmed to be false on social networks. Therefore, the extraction of posts was done on *fact-checking* articles, instead of direct extraction by *hashtags* or query searches on Facebook or Twitter itself. This resulted in a relatively small sample compared to other works that explore social networks. On the other hand, this ensured that the content analyzed in this work was necessarily associated with misinformation, as was desired.

Another limiting factor is that not all verified information contained in its body the original *post*, which began to be spread on social networks. Therefore, it is important the agencies to be careful to insert this kind of reference, since it allows the automation of analysis that can add value, especially for agencies that still carry out much of their work manually.

6 CONCLUSION

The COVID-19 pandemic impacted the whole world, through the economy, health systems, social and political relationships. Dealing with a situation as new as this in the information era is conducive to conspiracy theories and false information circulating across the web. This work endeavored to detect the topics that have been discussed to disseminate false information around the world, through social networks. For this, we relied on fact-checking agencies and showed that issues such as vaccine, cure, treatment, and deaths are being widely used to disseminate misinformation about the pandemic. The research also showed that these issues are being addressed by agencies, with the mission of reducing the damage caused by them, as it was possible to notice a drop in the number of posts made on social networks, as well as the share number of these posts.

The comparison between the fact-checking databases of Poynter and the Google Fact-Checking API showed that there is important work being done by the agencies, but that there is still a divergence in the data set. The Poynter Institute, for now, provides a more robust base concerning the volume of content, while the Google API has a more versatile and easy tool to automate processes, extremely suitable for programmers. The labels given by agencies are numerous, which can make interpretation and automation difficult for readers, as some labels are not very clear. It is difficult to differentiate, for example, between the classifications “partially true” and “partially false”. Also, some agencies did not provide a reference to the origin of the misinformation on the social network, which makes automation impossible in these cases. However, the agencies addressed pertinent topics, identified through the LDA, which meet what is being a reason for misinformation on social networks. Agencies have taken an average of 23 days to verify information posted on Twitter and approximately 6 days on Facebook. Still, there is a global movement in the fight against misinformation, as fact-checking agencies were identified in several countries.

The contributions brought by this work, unlike previous studies that analyzed public discourse during the pandemic, reveal findings directed towards the dissemination of false information and the activities of fact-checking agencies. This is a great challenge for public health, which must make efforts to fight not only the pandemic but also the “infodemia” [6]. The dissemination of false information on social networks can fuel widespread panic and cause harmful effects to the population, obscuring the evidence and making it difficult for professionals and public health systems to respond [1, 19] can help other researches around the dissemination of misinformation, including replicating the same methodology for other subjects, as the LDA can extract topics of high coherence, making it an appropriate tool to study infodemic phenomena in the context of public policy. It is important that this type of collaboration to be done so that not only the curve of COVID-19 cases decreases, but also the misinformation curve.

Social media data and topic modeling approaches facilitate the understanding of public discussions and concerns about the pandemic. Analyzing and understanding the real concerns of the public can increase the efficiency of the work carried out to combat the dissemination of false information.

As future works, it is intended to improve OCR techniques and other approaches to allow the exploration of a wider range of content contained in fact-checking stories, such as those circulated in WhatsApp, printed newspapers, blogs, and even those that circulated on social networks, but they do not have the redirect link. It is intended to carry out a study on the profile of those who share this type of information to better understand the phenomenon of misinformation.

ACKNOWLEDGMENTS

This work was supported by FAPEMIG (PPM-00253-18), CAPES and CNPq.

REFERENCES

[1] Alaa Abd-Alrazaq, Dari Alhuwail, Mowafa Househ, Mounir Hamdi, and Zubair Shah. 2020. Top Concerns of Tweeters During the COVID-19 Pandemic: Infoveillance Study. *J Med Internet Res* (22 4 2020). https://doi.org/10.2196/19016

[2] AFP. 2020. This video shows a mock funeral procession during an anti-citizenship law protest in India in 2019. Acessed 26/01/2021: https://factcheck.afp.com/video-shows-mock-funeral-procession-during-anti-citizenship-law-protest-india-2019.

[3] Guilherme Mossi Bento and Raul Teruel dos Santos. 2018. Avaliação de métodos de remoção de outliers e seus impactos na precisão dos métodos de interpolação. *1º Simpósio Mato-Grossense de Mecanização Agrícola e Agricultura de Precisão - SIMAP* 1 (2018). https://doi.org/10.1503/cmaj.109-5536

[4] David M. Blei and John D. Lafferty. 2009. Topic Models. *Chapman and Hall/CRC* (2009), 101–124.

[5] David A. Broniatowski, Amelia M. Jamison, SiHua Qi, Lulwah AlKulaib, Tao Chen, Adrian Benton, Sandra C. Quinn, and Mark Dredze. 2018. Weaponized Health Communication: Twitter Bots and Russian Trolls Amplify the Vaccine Debate. *American Journal of Public Health* 108, 10 (2018), 1378–1384. https://doi.org/10.2105/AJPH.2018.304567 arXiv:https://doi.org/10.2105/AJPH.2018.304567 PMID: 30138075.

[6] Matteo Cinelli, Walter Quattrociocchi, Alessandro Galeazzi, Carlo Michele Valensise, Emanuele Brugnoli, Ana Lucia Schmidt, Paola Zola, Fabiana Zollo, and Antonio Scala. 2020. The COVID-19 social media infodemic. *Scientific Reports* 10, 1 (06 Oct 2020), 16598. https://doi.org/10.1038/s41598-020-73510-5

[7] Chaolin Huang, Yeming Wang, Xingwang Li, Lili Ren, Jianping Zhao, Yi Hu, Li Zhang, Guohui Fan, Jiuyang Xu, Xiaoying Gu, Zhenshun Cheng, Ting Yu, Jiaan Xia, Yuan Wei, Wenjuan Wu, Xuelei Xie, Wen Yin, Hui Li, Min Liu, Yan Xiao, Hong Gao, Li Guo, Jungang Xie, Guangfa Wang, Rongmeng Jiang, Zhancheng Gao, Qi Jin, Jianwei Wang, and Bin Cao. 2020. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet* 10223 (15 2 2020), 497–506.

[8] Sophie E. Jordan, Sierra E. Hovet, Isaac Chun-Hai Fung, Hai Liang, King-Wa Fu, and andZion Tsz Ho Tse. 2019. Using Twitter for Public Health Surveillance from Monitoring and Prediction to Public Response. *Data* (1 4 2019). https://doi.org/10.3390/data4010006

[9] Heidi J. Larson. 2018. The biggest pandemic risk? Viral misinformation. 562 (16 10 2018). https://doi.org/10.1038/d41586-018-07034-4

[10] Stephan Lewandowsky, Ullrich K. H. Ecker, Colleen M. Seifert, Norbert Schwarz, and John Cook. 2012. Misinformation and Its Correction: Continued Influence and Successful Debiasing. *Psychological Science in the Public Interest* 13, 3 (2012), 106–131. https://doi.org/10.1177/1529100612451018 arXiv:https://doi.org/10.1177/1529100612451018 PMID: 26173286.

[11] Samuel Lima. 2021. Não há provas de que vídeo de mulher tremendo ao caminhar tenha relação com vacina da Pfizer. Acessed 06/03/2021: https://politica.estadao.com.br/blogs/estadao-verifica/nao-provas-de-que-video-de-mulher-tremendo-ao-caminhar-tenha-relacao-com-vacina-da-pfizer/.

[12] Maldita.es. 2020. Maldita.es – About Us. https://maldita.es/ Acessed 12/09/2021.

[13] Poynter. 2020. The CoronaVirusFacts/DatosCoronaVirus Alliance Database. (04 06 2020). https://www.poynter.org/ifcn-covid-19-misinformation/

[14] Michael Röder, Andreas Both, and Alexander Hinneburg. 2015. Exploring the Space of Topic Coherence Measures. In *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining (WSDM '15)*. Association for Computing Machinery, New York, NY, USA, 399–408. https://doi.org/10.1145/2684822.2685324

[15] Michael Röder, Andreas Both, and Alexander Hinneburg. 2015. Exploring the Space of Topic Coherence Measures. In *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining (WSDM '15)*. Association for Computing Machinery, New York, NY, USA, 399–408. https://doi.org/10.1145/2684822.2685324

[16] Mohsen Sayyadiharikandeh, Onur Varol, Kai-Cheng Yang, Alessandro Flammini, and Filippo Menczer. 2020. Detection of Novel Social Bots by Ensembles of Specialized Classifiers. *CoRR* abs/2006.06867 (2020). arXiv:2006.06867 https://arxiv.org/abs/2006.06867

[17] Edson C. Tandoc Jr, Zheng Wei Lim, and Richard Ling. 2018. Defining "Fake News". *Digital Journalism* 6, 2 (2018), 137–153. https://doi.org/10.1080/21670811.2017.1360143 arXiv:https://doi.org/10.1080/21670811.2017.1360143

[18] Andreas Vlachos and Sebastian Riedel. 2014. Fact Checking: Task definition and dataset construction. In *Proceedings of the ACL 2014 Workshop on Language Technologies and Computational Social Science*. Association for Computational Linguistics, Baltimore, MD, USA, 18–22. https://doi.org/10.3115/v1/W14-2508

[19] Lauren Vogel. 2017. Viral misinformation threatens public health. *CMAJ* 189, 50 (2017), E1567–E1567. https://doi.org/10.1503/cmaj.109-5536 arXiv:https://www.cmaj.ca/content/189/50/E1567.full.pdf

[20] WHO. 2020. Coronavirus disease (COVID-19) Situation Report – 139. *World Health Organization* (2020). Acessed 07/06/2020: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200607-covid-19-sitrep-139.pdf?sfvrs=79dc6d08$_2$.