Ref. Ares(2023)1854274 - 14/03/2023

sCience and human factOr for Resilient sociEty

Deliverable title:	Report about communication patters
Deliverable ID:	7.1
Document version:	1.0
Partner responsible:	CLB
Due date:	28/02/2023
Status:	Final

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

`

Document Control Sheet

Deliverable number	D 7.1
Deliverable name	Report about communication patters
Dissemination Level	PU
Call	H2020-SU-SEC-2020
Topic	SU-DRS01-2018-2019-2020 Human factors, and social, societal, and organisational aspects for disaster-resilient societies
Consortium Coordinator	UNISA
Edition	1

Authoring & Approval

Authors of the document	
Name/Beneficiary	Position/Title
Abraham Yosipof (CLB)	Task 7.1 leader
Carmit Rapaport (CLB)	Task 7.1 leader

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

`

Approved for submission - Representatives of beneficiaries involved in the project		
Name/Beneficiary	Position/Title	Date
Paolo Capuano	Project coordinator	12/03/2023
Raffaella Russo	Project coordination committee	09/03/2023
Dilanthi Amaratunga	Reviewer	27/2/2023
Richard Haigh	Reviewer	27/2/2023
Marie-Christine Bonnamour	Reviewer	03/03/2023

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

`

ABSTRACT

Social media platforms serve as an important source of information in times of emergency and disaster. The usage of social media provides a virtual place for people to gather, share their concerns, and ask for both physical and mental support. In fact, people use social media and other information platforms to learn about the event, its effect, and its causes, and might even elaborate on this information to enhance protective behaviour on their own as a preventive action. This information seeking behaviour is an expected response to the tragic, surprising event. However, it also reflects on a social process, which might shed light on the way in which disasters are perceived by the public, and react to them. In this research, we examined the communication patterns of European citizens on social media and information platforms over time and during and after disasters and emergencies. We performed four specific data analyses based on a unique social media database. The database is based on two sources: Twitter and Wikipedia. From Twitter we extracted tweets and the user's metadata related to specific case studies such as COVID-19 conspiracy theories, MonkeyPox and earthquakes. From Wikipedia we extracted the page traffic of seven case studies. The analysis included four stages: First, we explored the communication reactions to the conspiracy theories that emerged during COVID-19. Second, we examined a specific case of misinformation spread during the Monkeypox outbreak. Third, we focused on discussions of misinformation regarding earthquake "prediction". Finally, we analyzed disaster information seeking behaviour through the analysis of the traffic of disaster pages on Wikipedia. The results show the dynamics of the communication patterns and allow us to derive policy recommendations on measures to address misinformation and conspiracy theories on social media. The results indicated that people respond to global disasters – even if not in their country – which can lead to better preparedness planning and risk communication during disaster events.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

`

TABLE OF CONTENTS

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

LIST OF FIGURES

LIST OF TABLES

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

`

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

`

ACRONYMS

API: Application Programming Interface.

BERT: Bidirectional Encoder Representations from Transformers

CDC: Centers of Disease Control and Prevention

COVID-19: Coronavirus disease 2019

CT-BERT: Covid-Twitter Bidirectional Encoder Representations from Transformers

GMO: Genetically Modified Organism

KNN: K-Nearest Neighbors

NLP: Natural Language Processing

PADM: Protective Action Decision Model

RoBERTa: Robustly Optimized Bidirectional Encoder Representations from Transformers Pretraining Approach

SBERT: Sentence -Bidirectional Encoder Representations from Transformers

SVM: Support Vector Machine

URL : Uniform Resource Locator

WHO: World Health Organization

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

`

1. INTRODUCTION

Disasters' most significant characteristic is uncertainty. At the very first moments after the event had occurred, very little is known, even in the disaster area, about what is the cause of the event, how serious is the physical damage and most importantly, what (if any at all) is the cost in lives. Therefore, disaster management and search and rescue operations begin with creating a situation damage assessment based on victims' testimonies and calls for help. Only after operational forces, such as police officers, firefighters, and emergency medical services, arrive at the scene, the situation gets clearer, and search and rescue operations begin. However, in parallel with the disaster management efforts who physically save lives and help the victims, social media websites and applications might serve as important sources for information, especially for those who are not necessarily located in the affected area but are interested or emotionally effected by the event. This usage of social media provides a virtual place for people to gather, share their concerns, and ask for both physical and mental support. In fact, people use social media to learn about the event, its effect, its causes, and might even elaborate this information to enhance protective behaviour on their own as a preventive action. This information-seeking behaviour is an expected response to the tragic surprising event. However, it also reflects on a social process, which might shed light on the way in which disasters are perceived by the public, and react to them.

Disaster behaviour

The literature in the field of risk communication indicates on accessibility to information and its content as key factors leading to a change in behaviour [1]. Information search immediately after a disaster is critical for various functions including warning, situational awareness, getting instructions and finding social support.

A basic factor in emergency preparedness and response is information - in relation to risks on the one hand and the desired behaviour to reduce them, on the other hand [2-4]. Although countries invest significant amounts of resources in communicating with the public about emergency issues – including preparedness instructions, information about risks, actions of the emergency authorities, preparedness rates remain relatively low [5, 6].

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

In addition, perceptions about information on emergency preparedness were also found to be related to actual preparedness [7]. In a study that compared earthquake preparedness of citizens in New Zealand and Japan, it was found that, regardless of cultural background, the more the citizens perceive the emergency authorities as providing a response to their needs, the higher the level of trust they have in these organizations and the information they provide. As a result, the use of the information they provide increases in order to properly prepare for an earthquake.

The role of information and information seeking

Information is a critical element in disaster management, both for decision makers and the public. The emergency authorities use mass media to inform the public about potential risks and threats through alerts and warnings, instructions for action in various emergency situations, information about evacuation and post-disaster arrangements such as rehabilitation operations and return to normal.

The literature on risk communication highlights several approaches to the process of informing the public about potential risks. One is the importance of risk perception. The basic assumption holds that high levels of perceived risk will motivate taking preparedness actions, and therefore, there is a need to evaluate the perceptions of risk among the public or to motivate it through the transmission of messages [8, 9]. Models such as the "Information Likelihood Model" and the "Protective Action Decision Model" (PADM) show how information about threat changes attitudes, triggers decision-making and shapes action [10]. This is the basis for the need for appropriate dissemination of information about risks and desired behaviour [11]. Furthermore, Wood et al. [12] show that "communicating actionable risk" which includes explicit preparedness actions instead of information about the risks led to higher levels of preparedness. Therefore, the emphasis should be placed on the message itself, that is, the information conveyed to the public, and especially on clear instructions for implementation. The main factors that lead to taking action and changing behaviour in the communication process are: the goals of the communication itself (raising awareness, education, motivation to take action), the design of the content, the choice of the channel for the transmission of the message and the timing of the transmission of the message (for example, after

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

an event in a certain country or occasionally during routine) [1] and the source of the information [9, 12].

Information seeking includes different variables analyzed by Kyne and Donner [13]: information-seeking frequency, information-seeking behaviour and information sources. According to the authors, frequent information seeking, and authority information sources encourage protective response in compliance with authority's recommendations.

The motivation behind information seeking has been also explored. Zhu et al.[14] integrated the PADM and Heuristic–Systematic Model (HSM) and observe that perceived knowledge influences information seeking. The proposed model also provides insights on the impact of information seeking on systematic processing. According to the HSM, the receiver can process messages either systematically or heuristically. Systematic processing is cautious and reflective, while heuristic processing involves simplification and heuristics to quickly process the message.

The role played by the social media during and after emergencies.

The Internet - World Wide Web - serves as a significant tool for managing emergencies, in terms of the two-way and immediate interaction it enables between emergency authorities and citizens, between victims and citizens staying in the affected areas, first responders and local authorities and organizations [15, 16]. The use of websites or mobile applications to transmit information about preparedness and/ or response actions enables a wide and timely distribution of information with emergency authorities and potential recipients. Therefore, there is a need to use websites and mobile applications due to their visual advantages and their wide distribution to achieve motivation for action by cultural, cognitive and emotional adaptation to the recipient public [11]. The development of social media, and in particular social networks such as Facebook (Facebook.com) and Twitter (Twitter.com) and their growing usages in the context of emergencies, is becoming significant recently [17]. The first messages about emergencies that have occurred, for example, the attack on the Boston Marathon [18], came through Twitter. It was also found that social networks provide the individual with social and mental support through the communication with his family members, friends and acquaintances, as well as given the ability to find out relevant information [19].

The central characteristic of social media is the two-way communication and the interaction that produces information in a frequent and immediate manner [20, 21]. However, the information transmitted on social networks includes rumours,

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

`

partial information, information that has expired or is no longer relevant or even incorrect information [22]. Social media platforms, including Twitter, have a key role in the distribution of information regarding emergencies and disasters, including communicating reliable news, updates, and medical instructions. In addition, social media has been a fertile ground for the growth and distribution of misinformation and conspiracy theories.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

2. Methods

In this task, we aimed to examine the communication patterns on social media, namely Twitter and Wikipedia, in terms of changes over time, in reaction to disasters and emergencies, and in relation to conspiracy publishments. These platforms were selected due to several reasons, both platforms provide an API for downloading the data. In addition, Twitter is the only social media platform that provides an API for downloading the data for academic research. These two platforms represent both a social network (Twitter) and information seeking website (Wikipedia).

To meet the research objective, we created a unique database which is based on two sources: Twitter and Wikipedia. From Twitter we extracted tweets and the user's metadata related to specific case studies such as COVID-19 conspiracy theories, MonkeyPox and earthquakes. From Wikipedia we extracted the page traffic of seven case studies. The following sections explain in more details the datasets extracted for the Twitter Database and the Wikipedia traffic Database.

2.1 Twitter Database

The Twitter database includes tweets that were extracted from Twitter API and consists of three datasets according to the case study, namely: 1) The COVID-19 conspiracy theories tweet dataset, 2) The MonkeyPox tweet dataset, and 3) the Earthquakes prediction tweet dataset.

1) The COVID-19 Conspiracy dataset: a dataset of over 1.4M tweets and metadata of the users related to the eight COVID-19 conspiracy theories between 2020 and 2022. The eight COVID-19 conspiracy theories include the 5G, Big Pharma, Bill Gates, biological weapon, exaggeration, Film Your Hospital, Genetically Modified Organism (GMO), and the vaccines conspiracy (see table 1).

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

`

Table 1. Total number of tweets per conspiracy between 2020- 2022.

Conspiracy	Number of tweets
5G	419,324*
Big Pharma	187,805
Bill Gates	156,962
Weapon	232,832
Exaggeration	326,683
FilmYourHospital	7,152
GMO	18,445
Vaccines	72,961
Total	1,422,164

**The 5G includes 331,448 English tweets and 87,876 tweets in 19 European languages.*

2) The MonkeyPox dataset: A dataset with over 1.44M tweets related to the discussion on the monkeypox outbreak between May 2022 and August 2022.

3)The Earthquakes dataset: A dataset of over 80K tweets related to the discussion of misinformation on earthquakes prediction.

2.2 Wikipedia page Traffic Database

We collected the Wikipedia page traffic data for CORE six representative case studies in multiple languages, (the L'Aquila earthquake, Manchester Arena bombing, Aude River flooding, Visakhapatnam gas leak, Tōhoku earthquake and tsunami, COVID-19 pandemic) and 2021 European floods. For each case study, we extracted the traffic data of the case study's page in its official language, e.g., Italian for an earthquake in Italy, and usually English as well as a global language (see Table 1). For COVID-19, we used the data of five languages: Chinese, English, German, Italian and French. The database includes the traffic page data from the day the emergency started until January 2023.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

Table 2.The emergencies and disasters events

Event	Country	Year	Languages
L'Aquila earthquake	Italy	2009	Italian, English
Manchester Arena bombing	UK	2017	English
Aude river flooding	France	2018	French
Visakhapatnam gas leak	India	2020	English, Hindi
Tōhoku earthquake and tsunami	Japan	2011	Japanese, English
COVID-19 pandemic	Global	2019	English, German, French, Italian, Chinese
2021 European floods	Europe	2021	German, English

2.3 Data Analysis

To examine the communication patterns of European citizens on social media over time and during and after disasters and emergencies we investigated four case studies using data mining, machine learning and Natural Language Processing (NLP) tools.

First, we explored the communication reactions to the conspiracy theories which had emerged during COVID-19. Second, we examined a specific case of spread of misinformation in the case of Monkeypox outbreak in European countries. Third, we focused on discussions of misinformation regarding earthquake "prediction". Lastly, we analyzed disaster information seeking behaviour through the analysis of disaster pages on Wikipedia. In the next section we present the results of each of these examinations.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

3. Results

3.1 The COVID-19 Conspiracy Theories Communication Patterns Analysis:

The analysis and results of this section were published in two research papers [23, 24]. This section is structured as follows. Section 3.1.1 examines the conspiracy theories discussion on Twitter. Section 3.1.2 Analysis of the 5G conspiracy theory, regarding supporting or opposing the conspiracy. Section 3.1.3 Analysis of the communication patterns of the 5G conspiracy theory across Europe.

3.1.1 COVID-19 Conspiracy Theories Discussion on Twitter

The objective of this section is to analyze the changes in the discussion of different COVID-19 conspiracy theories throughout the pandemic on Twitter. The analysis includes tweets related to the eight most frequent COVID-19 conspiracy theories: the 5G, Big Pharma, Bill Gates, biological weapon, exaggeration, Film Your Hospital, Genetically Modified Organism (GMO), and the vaccines conspiracy. This section was published as a full paper [23].

To meet the research objective, we extracted from Twitter the COVID-19 Conspiracy dataset (see section 2.1) according to specific queries designed by experts in the field (see reference [23] for more details). The dataset includes 1.3M English tweets, see section 2.1, table 1 for the number of tweets in each conspiracy.

Figure 1 presents the evolvement of the discussion of conspiracy theories over time from January 2020 through December 2021 by month. The graphical analysis helps categorize conspiracy theories into four groups namely peak at the beginning of the pandemic, increase throughout the pandemic, Persistent theories and Multiple peaks.

Peak at the beginning of the pandemic. The first group includes the 5G and the FilmYourHospital conspiracy theories. We see a peak in April 2020, a sharp decline straight afterwards and then a gradual decline of the theory (see Figure 1). The FilmYourHospital conspiracy theory follows a similar pattern with a peak in April 2020 and quite a sharp decline afterwards. From Table 3, we can see a positive and significant correlation ($r = 0.134$, $p<0.01$) between the two conspiracies.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

`

Increase throughout the pandemic. The second group includes the Big Pharma and the vaccines-related conspiracy theories. The Big Pharma conspiracy fluctuated from March 2020 to July 2021 and then began to rise in frequency (see Fig. 1). The frequency of the vaccines-related tweets was constant and very low until October 2020. Then it went up and remained stable until June 2021 at a higher level, after which it began to rise sharply. From Table 3, we see a positive and significant correlation ($r = 0.593$, $p<0.001$) between these conspiracies time series.

Persistent theories. The third group is exaggeration and the Bill Gates conspiracy theories. The exaggeration conspiracy remained at a high level from March 2020 to January 2021, then declined until June 2021, then sharply reached previous high levels again in August 2021, and declined again (see Figure 1). However, it remained at a high level in November 2021. The conspiracy on Bill Gates appeared in 2021 and remained at a relatively high level with some fluctuations.

Multiple peaks. The fourth group are the GMO and the biological weapon conspiracies. Both reached a peak at the beginning of the pandemic and sharply declined afterwards to negligible levels (see Figure 1). However, the biological weapon theory peaked again in July 2021 with the peak being higher than in 2020. The GMO theory peaked in August 2021, though the second peak was lower. Afterwards, they both sharply returned to relatively low values.

We conduct a correlation (see Table 3) and a cross-correlation analysis to see how the conspiracy theories are related to each other. While the correlation between weapon conspiracy and vaccines-related conspiracies is significant and positive ($r=0.129$, $p<0.001$, see Table 3), the cross-correlation between the weapon tweet frequency at time t and the vaccines tweet frequency at time t+7 is much higher with a coefficient of 0.283. This result suggests that a higher tweet frequency of the weapon conspiracy at time t leads to a higher vaccine conspiracy tweet frequency seven days later (t+7). Bill Gates and weapons have no correlation between them (see Table 3), but the cross-correlation suggests that there is a positive correlation with a coefficient of 0.214 between the weapon at time t and the Bill Gates at time t+7. FilmYourHospital and GMO have a significant correlation coefficient of 0.176 (see Table 3), the cross-correlation suggests a higher coefficient of 0.593 between GMO at time t and FilmYourHospital two days later (t+2), meaning higher GMO tweet frequency leads to higher FilmYourHospital tweet frequency two d later. The 5G and Big Pharma have no significant correlation between them (see Table 3) but there is a cross-correlation with a coefficient of -0.247 between 5G at time t and Big Pharma at time t+5, and

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

cross-correlation of -0.22 between 5G at time t+4 and Big Pharma at time t, meaning that a higher tweet frequency of one of the conspiracies leads to a lower frequency in the second a few days later. The 5G and vaccines conspiracies are negatively significantly correlated (r = -0.212, p<0.001, see Table 3). The cross-correlation suggests a coefficient of 0.25 at lag +4 and lag -4 meaning that a higher vaccines tweet frequency at time t leads to a lower 5G tweet frequency 4 days later, and vice versa (higher 5G tweet frequency leads to lower vaccines tweet frequency 4 days later).

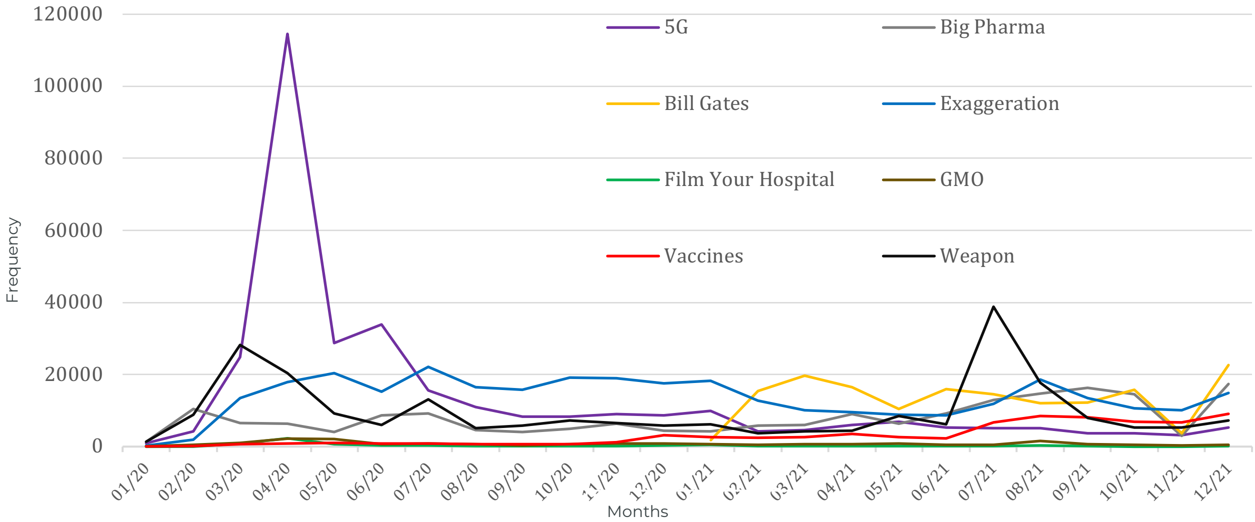


Figure 7. Frequency of tweets per month, grouped by theory.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

Table 3. Correlation matrix between the conspiracy theories tweets frequency time series.

	5G	Bill Gates	Big Pharma	Weapon	Exaggeration	FilmYour Hospital	GMO	Vaccines
5G	1							
Bill Gates	0.293***	1						
Big Pharma	-0.032	0.076	1					
Weapon	0.152***	0.004	0.121**	1				
Exaggeration	0.164***	0.095	0.056	0.087*	1			
FilmYourHospital	0.134**	-0.058	-0.090*	-0.022	0.182***	1		
GMO	0.080*	0.043	0.019	0.032	0.099*	0.176***	1	
Vaccines	-0.212***	0.128*	0.593***	0.129***	-0.028	-0.172***	-0.027	1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

3.1.2 Analysis of COVID-19 5G Conspiracy Theory Tweets

The COVID-19 5G Conspiracy Theory Tweets analysis is the first attempt to classify tweets related to the COVID-19 5G conspiracy theory into supporters and opponents of the conspiracy theory and analyzing the results. The analysis was published as a full paper[24].

A study of conspiracy theory tweets presents several challenges, such as the collection of enough relevant data over a long time period for the classifiers, but at the same time not collecting too much irrelevant data. Another challenge is to embed the semantic meaning of the tweets in vectors of features, labeling, and classifying the data. Finally, an analysis of the tweets is needed in order to gain insights about how the conspiracy evolves.

A key task in analyzing conspiracy or misinformation tweets is to label and classify the tweets. A set of embedding features and labels are needed for the training of a classifier. The Bidirectional Encoder Representations from Transformers (BERT) provides superior results for different NLP tasks, including word embedding [25, 26]. Micallef et al. [27] used BERT embeddings to investigate and counter

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

misinformation in tweets related to COVID-19 over a period of five months. In that work, the authors trained a classifier and classified a dataset of 150K COVID-19 related tweets using BERT embeddings and analyzed the results.

Different metadata and characteristics of tweets and their authors were proven to be useful for classification tasks and were previously used to enhance classification models. Beskow and Carley [28] used the number of users that follow the author and the number of users the author follows as an indication of whether the author is a robot or not. O'Donovan et al. [29] and Gupta et al. [30] found that metadata of tweets, such as URLs, mentions, retweets and tweet length may serve as indicators for credibility.

Our objective is to develop a workflow to collect and classify tweets as supporting, opposing, or neutral of a conspiracy theory, in order to analyze the discussion on COVID-19 5G conspiracy theories on Twitter. We therefore suggest the following workflow as described in Figure 2:

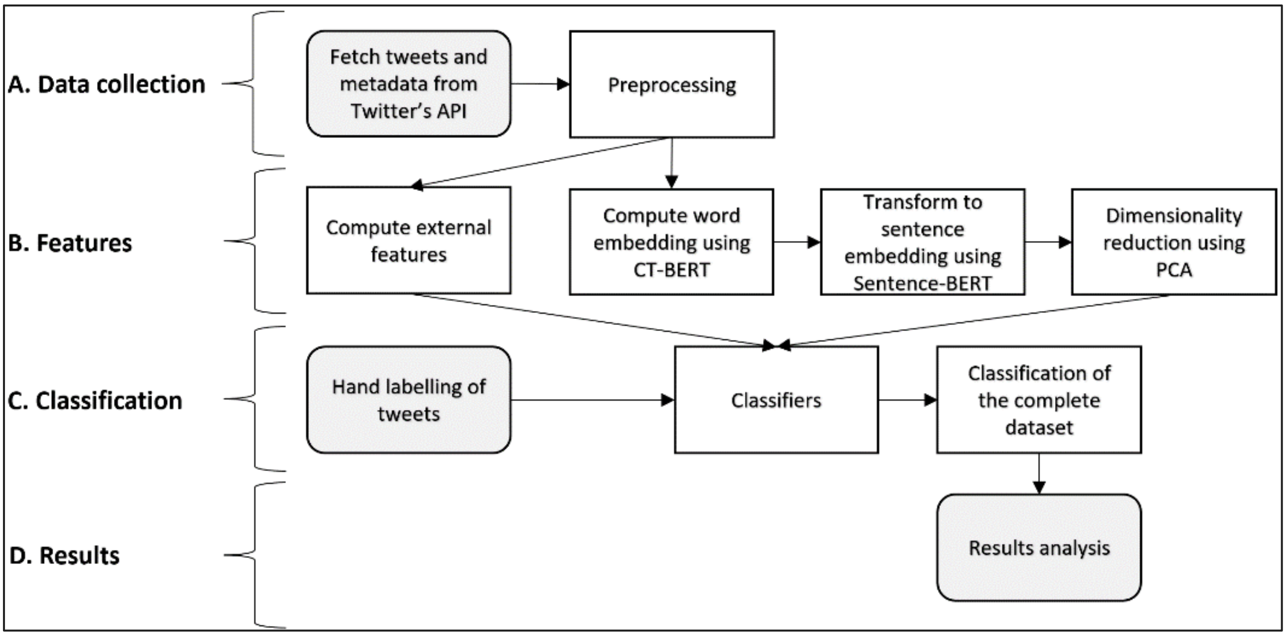


Figure 2. COVID-19 conspiracy theory analysis workflow.

(A) Data collection: collect tweets related to the COVID-19 5G conspiracy theory over a period of two years; (B) Features: compute a set of word embedding features using CT-BERT, transform to sentence embeddings using SBERT, and compute a set of external features; (C) Classification: develop a classifier based on

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

the sentence embeddings and external features to classify tweets as supporting or opposing of the conspiracy; and (D) Results analysis of the classified dataset.

(A) Dataset. We collected 331,448 English tweets related to the discussion on the COVID-19 5G conspiracy theory and metadata of the users that posted the tweets, over a period of two years from January 1, 2020, shortly after the pandemic emerged, to December 31, 2021 (see section 2.1).

We hand labeled 4,291 tweets as belonging to one of three categories: 2,147 supporters, 676 opponents, and 1,468 neutral/irrelevant of the conspiracy theory. Table 4 displays example tweets and their manual classification.

Table 4. Examples of 5G tweets and their manual classification.

Tweet	**Classification**
"5G is killing people. Covid was a cover. Now the jab is a cover."	Supporter
"You can get COVID through 5G."	Supporter
"#lie: 5G mobile networks DO NOT spread COVID-19. #coronavirus"	Opponent
"I can't believe there's people out there that actually think 5G causes covid"	Opponent
"Check out: Ericsson Revival Rides 5G, R&D "One critical factor, COVID, has been present in almost all discussions during the year," CEO Börje Ekholm said."	Neutral/ Irrelevant
"5G, AI, cybersecurity and renewable energy set for investment boost under EU coronavirus recovery plan"	Neutral/ Irrelevant

(B) Features: we computed two sets of features for each tweet. First, the sentence embedding using CT-BERT and SBERT. Second, external features that we computed or extracted from the tweets, authors, or their metadata.

Embedding Features: The Bidirectional Encoder Representations from Transformers (BERT) [31] is Natural Language Processing (NLP) model. BERT provides state of the art performance on data that is relevant to specific NLP tasks

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

such as representing the semantic meaning of the tweets in vectors of features [32]. Covid-Twitter-BERT (CT-BERT) is a model based on BERT Large, that was pre-trained and fine-tuned on a corpus of 160M tweets about the coronavirus [33]. The data on which CT-BERT was pre-trained and fine-tuned fits the scope of this work.

Sentence-BERT (SBERT) is a modification of the pre-trained BERT network that uses siamese and triplet network structures on top of the BERT model and fine-tuned based on high quality sentence interface data to learn more sentence level information [34]. SBERT transforms the CT-BERT word embedding into single sentence embedding with 1,024 features.

In this work, we computed the embedding of each tweet in the dataset using the CT-BERT model for the word embedding and SBERT to transform it to sentence embedding. To avoid overfitting of the classification models, we used Principal Component Analysis (PCA) to reduce the number of features. We selected the new principal components with eigenvalues greater than 0.1, which explain 82% of the variance of the 1,024 original features. The final vector of each sentence embedding consists of 211 features.

External Features: In addition to the sentence embedding, we computed and extracted six additional features from the metadata of tweets and their authors. Table 5 presents descriptive statistics of the external features. The six features include the sentiment score of each tweet as computed by VADER (Table 5, VADER), a parsimonious rule-based model for general sentiment analysis [35]. The sentiment score of each tweet is between -1 and 1, for negative and positive sentiment, respectively. We also used features based on the metadata of each tweet and author. These include the total number of tweets the author has posted (Table 5, Tweets), the average VADER of all of the conspiracy tweets we collected by the user (Table 5, Author avg. VADER), the presence of a URL in the tweet (Table 5, URL), the author's number of followers (Table 5, Followers), and how many users they are following (Table 5, Following). The external features were standardized using Z-score to prevent bias of the models toward high numbers.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

Table 5. Descriptive statistics of the external features.

Feature	Min	Max	Mean	Std.
Tweets	1	7,613,045	47,132.8	139,928.9
Author avg. VADER	-1	1	-0.12	0.39
Followers	0	55,462,408	28,056.6	531,498.2
Following	0	594,127	1,865.2	9,439.1
VADER	-1	1	-0.12	0.46
	Min	Max	# 0	# 1
URL	0	1	168,435	163,013

(C) Classification: We used five classification methods to classify each tweet as supporting, opposing, or neutral/irrelevant of the conspiracy theory.

The five methods are XGBoost with a learning rate of 0.3, Random Forest (RF) with 1,000 trees, Support Vector Machine (SVM) with a linear kernel function, K-Nearest Neighbors (KNN) with nine nearest neighbors, and Naïve Bayes. We further used voting ensemble learning to combine the results of all five methods.

In order to evaluate the performance of each model over the different sets of features, each classifier was trained on the embedding-based features (211 features), on the external features (6 features), and on the embedding and the external features together (217 features).

We evaluated the models using stratified 10-folds cross validation on the hand labeled tweets. The performance of each model was evaluated using the weighted F1, precision, and recall scores.

(D) Results and Analysis: The hand labeled training set was assigned to each of the classifiers as detailed in the classification section. Table 6 presents the classification performance for each model with the corresponding standard deviation when using the embedding features, the external features, and both sets of features together.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

Table 6 shows that the Voting Ensemble model provided the best results using the embedding features and the external features together, with weighted F1, precision and recall scores of 0.904, 0.907, and 0.903, respectively. The Voting Ensemble also performed well when only the embedding features were used, but with lower results compared to the combination of the embedding features and the external features. The models that were trained on the external features alone, provided poorer results compared to embedding features alone and both sets of features together.

Voting Ensemble with both sets of features is therefore the best method for the classification of the complete dataset. Following these results, we applied the Voting Ensemble model with both sets of features to classify the unlabeled dataset.

Table 6. Classification performance metrics.

Model	F1	Precision	Recall
Embedding + External Features			
XGBoost	0.895±0.02	0.896±0.02	0.894±0.02
Random Forest	0.891±0.02	0.893±0.02	0.891±0.02
KNN	0.878±0.02	0.889±0.01	0.876±0.02
SVM	0.877±0.02	0.877±0.02	0.877±0.02
Naive Bayes	0.742±0.02	0.791±0.02	0.736±0.02
Voting Ensemble	**0.904**±0.02	**0.907**±0.02	**0.903**±0.02
Embedding Features			
XGBoost	0.890±0.02	0.892±0.02	0.890±0.02
Random Forest	0.888±0.02	0.890±0.02	0.889±0.02
KNN	0.879±0.02	0.890±0.02	0.877±0.02
SVM	0.881±0.02	0.882±0.02	0.881±0.02
Naive Bayes	0.737±0.02	0.787±0.01	0.732±0.02

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

Voting Ensemble	**0.902**±0.02	**0.904**±0.02	**0.901**±0.02
External Features			
XGBoost	**0.597**±0.02	**0.595**±0.02	**0.614**±0.02
Random Forest	0.594±0.02	0.591±0.02	0.608±0.02
KNN	0.531±0.03	0.527±0.03	0.557±0.03
SVM	0.515±0.03	0.534±0.03	0.565±0.02
Naive Bayes	0.516±0.02	0.506±0.04	0.561±0.02
Voting Ensemble	0.564±0.02	0.585±0.03	0.599±0.02

Table 7 presents the results of the classification of the unlabeled dataset. The results show that 64,080 of the tweets support the conspiracy theory, 108,175 oppose the conspiracy theory, and 159,193 are neutral or irrelevant. Noticeably, there are 69% more tweets opposing the conspiracy theory than there are tweets supporting it. Examining the number of tweets per user in each category reveals that for each user supporting the conspiracy theory, there are 2.53 users opposing it. On the other hand, supporters of the conspiracy theory posted significantly more tweets per user, with an average of 1.82 tweets while opponents posted only 1.22 tweets on average.

Table 7. The number of tweets and users in each category.

Category	# Tweets	# Users	Tweets /User
Supporters	64,080	35,169	1.822
Opponents	108,175	89,030	1.215
Neutral/Irrelevant	159,193	100,298	1.587

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

Following the classification of the complete dataset, we analyzed the textual properties and frequency over time of the supporters and opponents of the conspiracy. We begin with examining the distribution of the classification categories by the sentiment of their tweets. Figure 3 presents a histogram with the density of tweets per classification category by their sentiment scores. The results show that the majority of the tweets in both categories have had a neutral sentiment score between 0 and 0.1. It can further be noticed that the frequency is denser with negative sentiments for both opponents and supporters of the conspiracy.

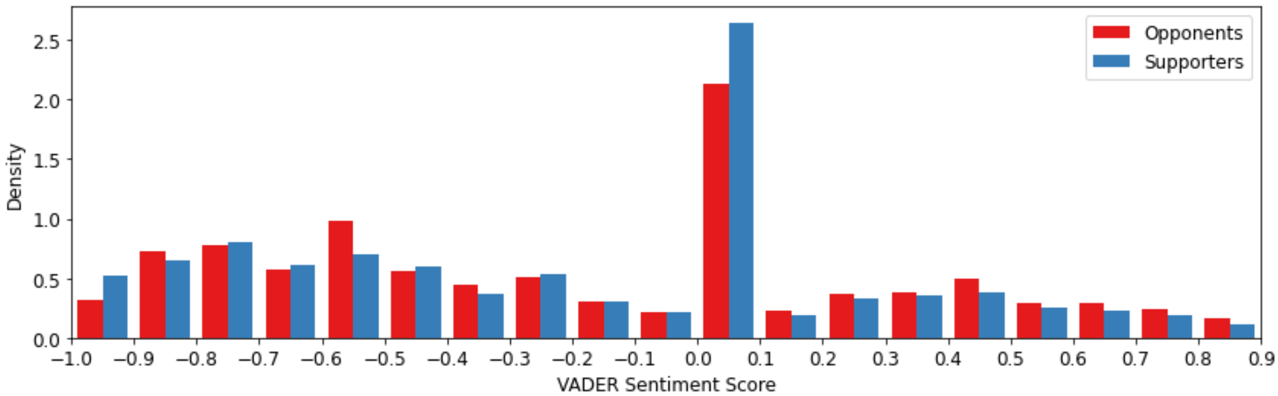


Figure 3. Density histogram of tweets' sentiment scores per classification category.

Figure 4 presents the monthly frequency of tweets per classification category. The conspiracy theory is noticeably declining over time in both categories. The monthly frequency of tweets per category emphasizes that when the conspiracy theory emerged around February 2020, there were substantially more tweets from supporters of the conspiracy theory. However, the balance changed in April 2020 when opponents have possibly started to confront the conspiracy theory. Since April 2020, there have been more tweets from opponents than supporters, which may indicate users have been educated.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

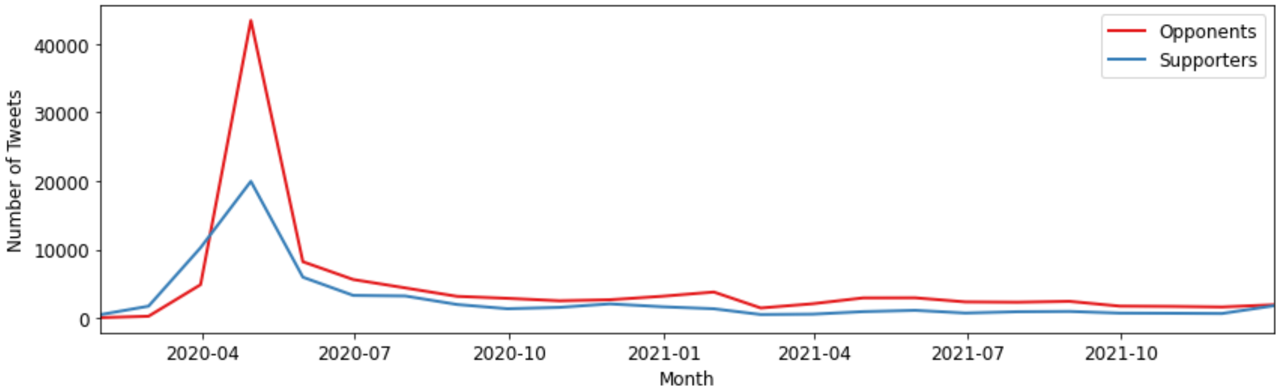


Figure 4. The monthly frequency of tweets per category for the 5G COVID-19 conspiracy theory.

The presence of a URL in a tweet, which is also one of the external features used in the classification process, shows how supporters and opponents use online resources to reinforce their opinions.

Figure 5 shows the number of tweets with and without a URL in each category. The analysis shows that only 28% of the tweets posted by opponents of the conspiracy theory linked a URL.

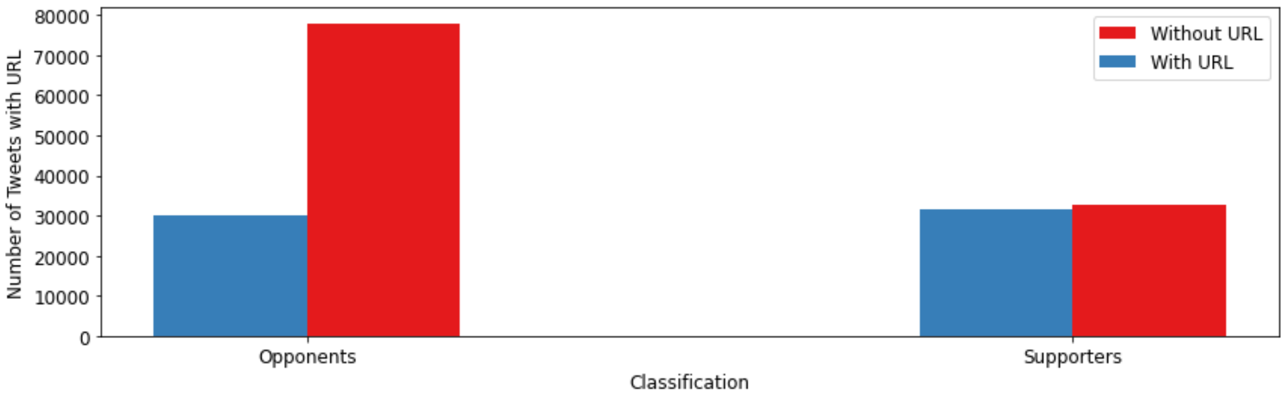


Figure 5. Number of tweets with and without a URL in each category.

This finding that opponents of the conspiracy theory include less evidence to refute the conspiracy, is not surprising and lines with Micallef et al. [27]. On the other hand, 49% of the tweets supporting the conspiracy theory did include a URL, likely to serve as evidence or to support the content of the tweet.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

3.1.3 Analysis of COVID-19 5G Conspiracy Theory Tweets Across Europe

The objective of this section is to analyze the differences on the discussion of the 5G COVID-19 conspiracy theory as a case study across 20 European languages.

We collected 419,324 tweets in 20 European languages related to the discussion on the COVID-19 5G conspiracy theory over a period of two years from January 1, 2020, to December 31, 2021. See

Table 8 presents the list of the languages and the number of tweets collected in each language.

Table 8. The number of 5G conspiracy tweets in each language.

Language	Tweets	Language	Tweets
English (en)	331,448	Lithuanian (lt)	949
French (fr)	23,292	Danish (da)	774
Portuguese (pt)	15,966	Romanian (ro)	607
Italian (it)	14,096	Slovenian (sl)	400
Dutch (nl)	12,884	Bulgarian (bg)	383
German (de)	10,496	Hebrew (iw)	355
Polish (pl)	3,526	Latvian (lv)	324
Finnish (fi)	1,281	Estonian (et)	232
Czech (cs)	975	Spanish (es)	198
Greek (el)	961	Hungarian (hu)	177

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

`

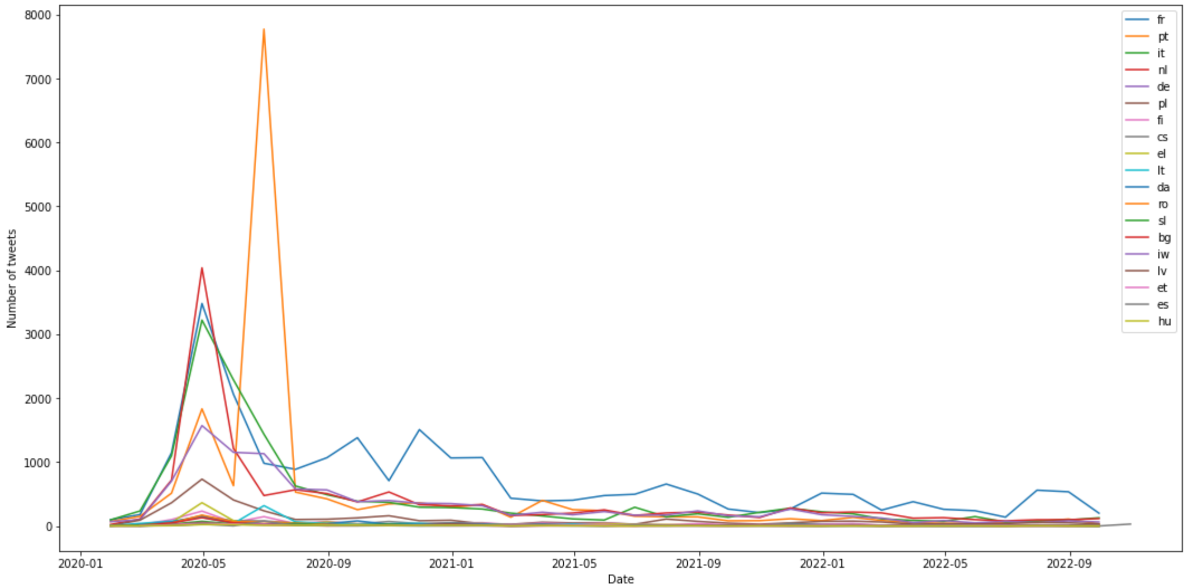


Figure 6. The monthly 5G conspiracy theory tweet frequency for each one of the European languages.

Figure 6 presents the monthly tweet frequency for each one of the European languages. Most European languages (beside English) are correlated at time t between each other, however no substantial cross-correlation was found at other lags.

The English tweets on the other hand, have shown cross correlation with European languages. Cross correlation was found between Danish at time *t+1* and English at time *t* (r=0.626), Greek at time t+1 and English at time t (r=0.622), and Italian at time t+1 and English at time t (r=812). These results suggest that higher 5G tweet frequency in English leads to higher 5G tweet frequency in Danish, Greek, and Italian a day later.

The results also show a cross correlation between Dutch at time t+7 and English at time t (r=0.814), which means that Dutch follows English after a week, and between French at t-5 and Dutch at t (r=0.683), which means Dutch is following French after 5 days. Accordingly, English at time t and French at time t+2 are cross correlated (r=0.686). meaning that higher 5G tweets frequency in English leads to higher number of 5G tweets in French a day later.

Spanish and Portuguese are highly correlated at time t (r=0.877) but no substantial cross correlation was found. This finding is logical as Spain is Portugal's only land border. Additionally, the 5G discussion Peaks were found in the time series in all European languages during April 2020. The Portuguese, Lithuanian and Hebrew time series also peaked again at June 2020.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

`

3.2. Mining the Discussion of Monkeypox Misinformation on Twitter

The objective of this section is to analyze the discussion related to monkeypox on Twitter and differentiate between tweets that spread and counter misinformation. Understanding the evolvement and behaviour of misinformation on social media enables **a better and faster reaction** to misinformation and conspiracy theories in the future. This study achieves a better understanding of the life cycle of misinformation and conspiracy theories using NLP and inspection of behavioural patterns.

We address these challenges and make the following contributions. We collected 1,440,475 tweets that are relevant to the discussion on monkeypox from 505,163 users on Twitter. We manually labelled 3,218 tweets into three categories, namely misinformation, counter misinformation, and neutral. We fine-tuned a Robustly Optimized BERT Pretraining Approach (RoBERTa) model for the classification task and compared its performance to several other machine learning classifiers. We analyzed the classified dataset to find and compare behavioural patterns in the data. Finally, we offer policy suggestions to reduce unwanted behaviour of misinformation and conspiracy spreading, and support wanted behaviour that counters misinformation and conspiratory narratives.

While in section 3.1.2 we used the CT-BERT a pretrained language BERT model that was pre-trained and fine-tuned on tweets about the coronavirus [33], in this section we needed to trained a model. Multiple variations of BERT with different strengths and weaknesses are available for a variety of tasks. RoBERTa was pretrained using different design decisions than BERT that improve the performance and state of the art results on different datasets [36]. Multiple studies and experiments evaluated the performance of RoBERTa and found that it provides better results than BERT [37-39]. In this section, RoBERTa model was chosen for the word embedding and classification tasks.

Dataset: We collected the dataset using Twitter's academic research API. The search query includes all tweets in English that contain the term "monkeypox" between May 1, 2022 and August 24, 2022, and excludes retweets. The query is simple yet very effective in filtering tweets related to the discussion on monkeypox, and the very wide search query enables us to collect a large amount of data without much noise.

We applied a preprocessing methodology similar to Nugen et al.,[40]. For performance optimization, we limited the dataset to tweets that were 350

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

characters or shorter after preprocessing. After preprocessing, the final dataset consists of 1,440,475 tweets related to discussion on monkeypox that were posted by 505,163 different users.

In order to train a classifier to classify the full dataset into three categories, namely misinformation, counter-misinformation, and neutral, we hand-labeled 3,218 tweets based on and according to facts provided by the World Health Organization [41]. Table 9 presents the number and examples of tweets in each category.

Table 9. Number and examples of tweets in each category.

Category	# Tweets	Tweet
Misinformation	1,090	*"The monkeypox travel with Pfizer vaccines. It can cover big distances quickly. Stop the vaccine traveling and that will stop the "monkeypox" virus also….."*
Counter-Misinformation	739	*"Monkeypox is a potentially serious disease caused by infection with the monkeypox virus. Anyone can get monkeypox and it's important for everyone to take precautions to stop the spread."*
Neutral	1,389	*"#India confirms Asia's first #monkeypox death."*

Classification: We fine-tuned a RoBERTa model, calculated the word embedding resulting in 768 features for each tweet and added a classification layer. We evaluated the model by stratified 5-fold cross-validation, each 20% test was further split to 10% test and 10% validation for the fine-tuning. We fine-tuned the model for 10 epochs with a dropout of 0.2, weight decay of 0.01, learning rate of 2e-5, and batch size of 16. The model converged at epoch 5 with an average validation loss of 0.631 and an average F1 score of 0.77 on the test sets. The performance results shows that the fine-tuned RoBERTa model achieves an average F1 score of 0.767, average precision of 0.774, and an average recall of 0.774.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

Results Analysis and Discussion: We assigned the unlabeled dataset to the RoBERTa model. The results of the classifier can be seen in Table 10. The results show that 180,259 of the tweets spread or support misinformation related to the monkeypox virus. 152,522 of the tweets counter misinformation related to the monkeypox virus, and 1,107,694 of the tweets are neutral to the discussion.

The classifier results indicate that most of the monkeypox tweets are neutral to the discussion on misinformation. This methodology allows us to quantify the scale of the discussion on misinformation in relation to the general discussion on monkeypox. The results show that merely 30% of the tweets on monkeypox discuss misinformation.

Table 10. The number of tweets and users in each category.

Category	# Tweets	# Users	Tweets/User
Spread misinformation	180,259	112,765	1.599
Counter misinformation	152,522	94,748	1.610
Neutral	1,107,694	401,458	2.759

Analyses of the misinformation and counter-misinformation categories indicates that for each tweet that spreads misinformation there are only 0.85 tweets that counter it. For each user that spreads misinformation there are 0.84 users that counter them. However, users that counter the conspiracy theories tweet slightly more per user than users that spread misinformation.

We examined the behavioural patterns in and between the misinformation and counter-misinformation categories. Figure 7 presents the weekly frequency of tweets in each group.

Two major peaks, likely related to the epidemiological evolvement of the virus, are observed. The first is during May 2022, when according to data from the WHO, the number of new cases reported globally started growing exponentially [42]. The second is in July and August 2022, during a peak of new confirmed cases.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

Tweets that spread misinformation have dominated the conversation since the beginning of the outbreak. However, a shift in dominance took place at the beginning of the second peak, possibly marking the last cycle of misinformation.

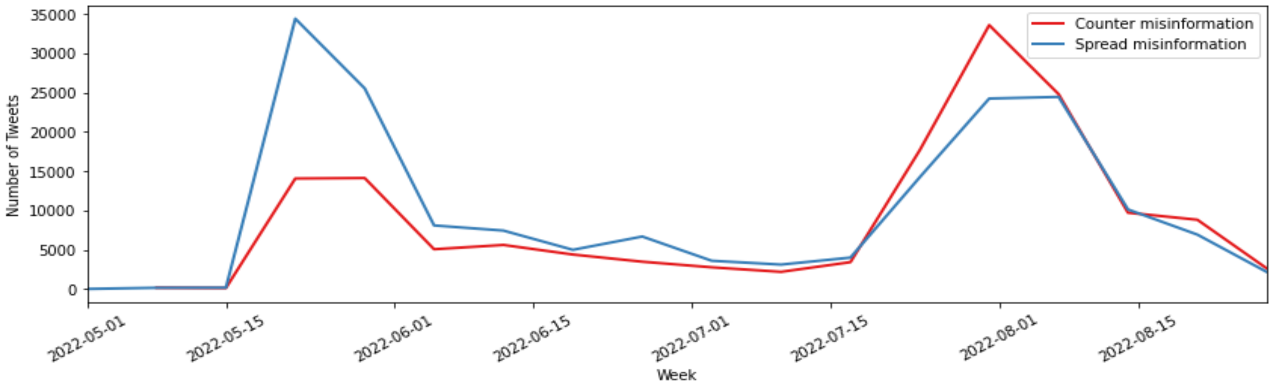


Figure 7. Weekly frequency time series of tweets spreading misinformation and tweets countering misinformation. Blue line represents the spreading tweets and red line represents the countering tweets.

We analyzed the users that participated in the conversation over time. Figure 7 presents the number of unique users that participated in the discussion each week, as well as the number of new unique users who participated in the discussion for the first time. The results show that the discussion is mostly driven by users that participate in the discussion for the first time. This behaviour indicates that the interest in the discussion on misinformation related to the monkeypox virus is authentic and not artificially created by a small number of users.

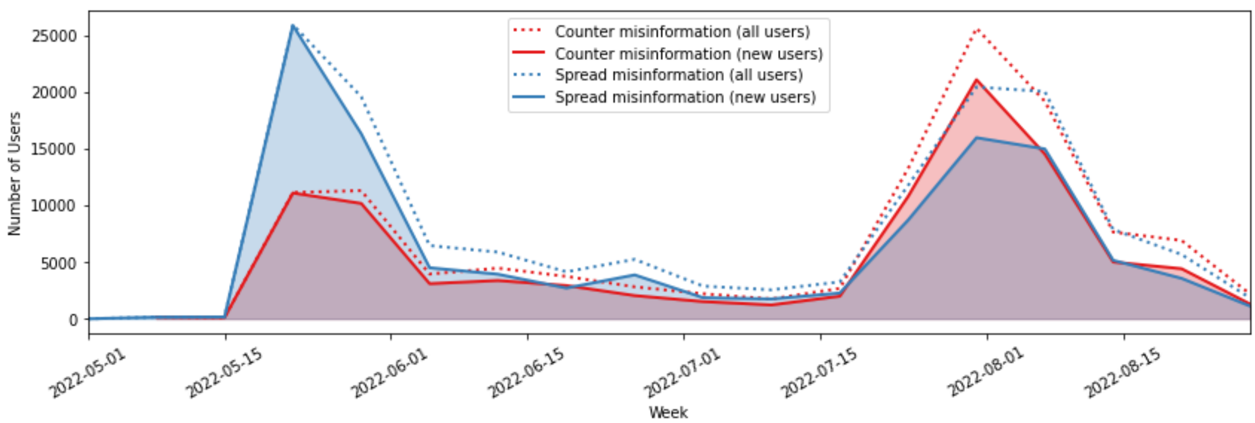


Figure 8. The weekly number of unique participating users and new unique users.

To investigate patterns in the behaviour of users, we analyzed the domains that were linked in the misinformation and counter-misinformation categories.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

Domains that do not provide any value to the analysis, such as social media websites and URL-shortening services were ignored.
Both the spreading and countering categories started by referencing news agencies based in the United Kingdom when the outbreak started. The reason for this behaviour is likely because the first cases of the outbreak were reported in the United Kingdom and Ireland [42]. The categories have quickly diverged and started referencing different websites from that point onwards.

Figure 9 and Figure 10 show the top 10 most referenced domains in each category and their percentage in the respective category. The most referenced domains in the category that counters misinformation are of authorities such the U.S. Centers of Disease Control and Prevention (CDC), globally acknowledged entities such as the WHO, and established news agencies such as the NBC and the New York Times.

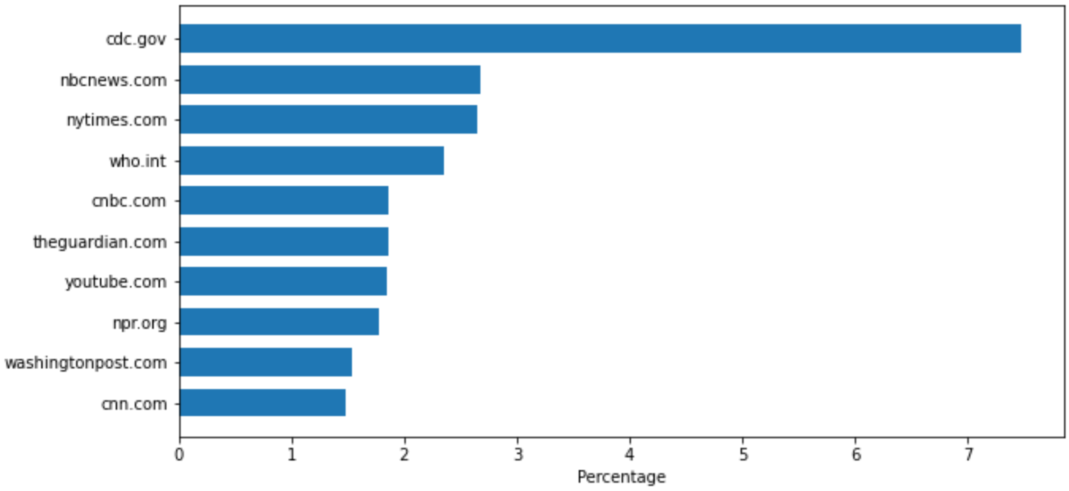


Figure 9. Top 10 most referenced domains in the not-misinformation category and their percentage in the category.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

`

On the other hand, the domains referenced in the category that spreads misinformation are mostly of websites that allow users to upload and publish their own content. YouTube leads the list, followed by other platforms that are often associated with extreme free speech and conspiracy theories.

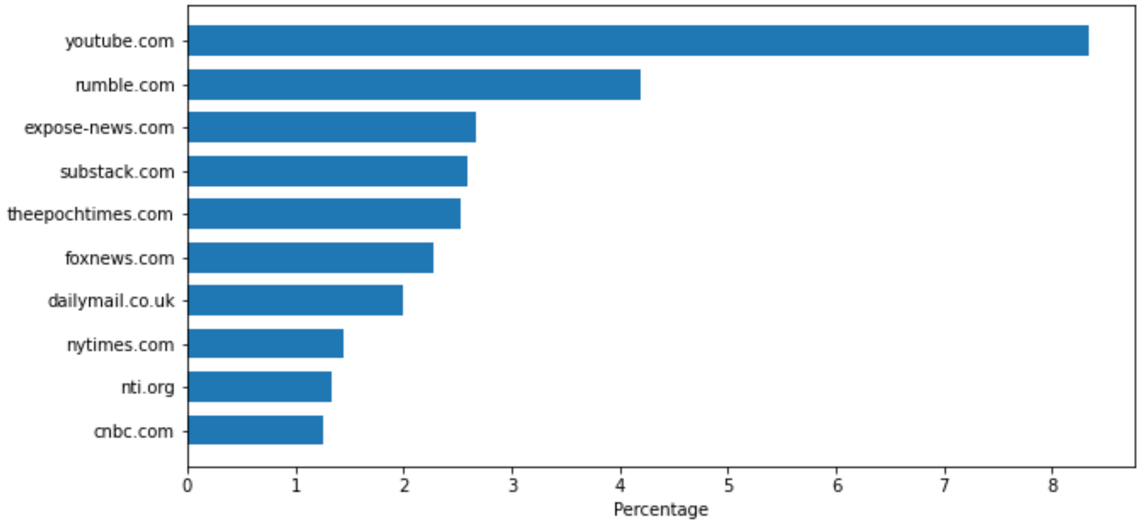


Figure 10. Top 10 most referenced domains in the misinformation category and their percentage in the category.

These findings are in line with the findings of Micallef et al. [27] that investigated misinformation on COVID-19 and found that YouTube is the most frequently referenced website in both misinformation and counter-misinformation.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

3.3. Earthquake "predictions" on Twitter

One of the issues of misinformation on Twitter is earthquake prediction. Currently, earthquakes cannot be predicted, meaning that the exact location, time, and magnitude of the next large event cannot be specified. What scientists are able to do is to provide a forecast, thus estimate the probability of certain magnitude earthquakes to occur in each space-time magnitude domain [43]. we conducted a study to analyze the dynamics and patterns of earthquake prediction statements on Twitter. To the best of our knowledge, this is the first study to collect, classify, and analyze all English tweets that discuss misinformation related to earthquake predictions, over a period of two years.

we collected 82,129 tweets related to the subject of earthquake misinformation, and metadata of the users that posted the tweets, over a period of two years from March 1, 2020, to March 31, 2022.

In order to analyse the tweets we classified them into three categories. We hand-labelled 3,584 tweets into three categories: misinformation, not-misinformation, and irrelevant tweets by using professional seismologists and according of the Earthquake Misinformation Communication Guide [44].

The misinformation category includes all tweets that claimed to be able to predict a future earthquake, namely the precise location and time of it, according to the state of the art summarized in the recently published Communication Guide [44]. The not-misinformation category includes general notifications about current earthquakes, as well as tweets that clarify that earthquake predictions are not possible. All other tweets, such as tweets unrelated to earthquakes, discussion of secondary hazards of earthquakes such as volcanic eruptions, were classified as irrelevant. The final labelled dataset consists of 3,584 tweets 698 misinformation tweets, 1,328 not misinformation tweets and 1,558 irrelevant tweets. We used the hand-labelled dataset to train a machine learning classifier to classify the full dataset into the three categories.

Classification: We fine-tuned RoBERTa-base model using the label dataset five times for 10 epochs with a dropout of 0.2, weight decay of 0.01, learning rate of 1e-5, and a batch size of 16. We evaluated the models using stratified 5-folds cross validation on the labeled tweets. We split the test set into 10% for the validation set and 10% for the test set. The performance of each model was evaluated using the weighted F1, precision, and recall scores. To reliably evaluate the performance of the fine-tuned model, each checkpoint was restored and tested against the corresponding test set. The model converged at epoch 5 with an average

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

evaluation loss of 0.456 and an F1 score of 0.836±0.017 with precision and recall of 0.846±0.013 and 0.846±0.013, respectively, on the test set. The fine-tuned RoBERTa model was used for the classification of the complete dataset into one of three categories (see Table *11*)

Results and discussion: we analyzed the temporal dynamics characteristics of earthquakes tweets over time. In total, 82,129 tweets were considered for the analysis. Table 11. Number of tweets and users, and tweets per users for each category provides the classification results were 39,266 were found to be irrelevant, thus not referring to earthquakes or information about secondary hazards triggered by earthquakes. This relatively high number of irrelevant tweets is a consequence of the chosen search criteria that provides wide coverage of the discussion.

Table 11. Number of tweets and users, and tweets per users for each category

Category	# of Tweets	# of Users	Tweets/User
Misinformation	8,542	3, 786	2.26
Not misinformation	34,321	4,048	8.48
Irrelevant	39,266	28,002	1.40
Total	*82,129*	*35,836*	*2.29*

Further, 8,542 tweets contained misinformation stating that earthquakes can be predicted. The range of arguments was broad: i) self-announced experts with their own websites; ii) people claiming that the planet constellation predicts the next big earthquake; iii) individuals stating that a religious leader (e.g., God) is angry and will soon trigger an earthquake; iv) people saying that animals predicted the earthquake and v) there is also an on-going discussion (within the scientific community) that progress in machine learning will allow the community to predict future earthquakes.

The remaining 34,321 tweets – classified as not-misinformation - were primarily general earthquake notifications from official sources providing information about the location, time and affected area of an event that occurred. A small part

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

of these tweets were specific tweets clarifying that earthquakes cannot be predicted.

Overall, there thus are substantially more not-misinformation tweets than misinformation tweets, which means that accurate and reliable information dominates the twitter environment. Further, there are also more users in the not-misinformation group and these users tweet more (8.48 tweets/user) than the users in the misinformation group (2.26 tweets/user); see Table 11. Number of tweets and users, and tweets per users for each category

We further analyzed the frequency over time of both 'misinformation' and 'not-misinformation' tweets. Figure 11. Daily frequency of tweets for earthquake discussion. The daily peaks often correlate between the two groups, showing that during an earthquake sequence or after an event the spread of predictions increased.

For instance, on February 10, 2021 (Figure 11, annotations 4) show that high magnitude earthquake triggering many aftershocks with relatively high magnitudes receive internationally high intention. On the February 10, 2021, when the Mw 7.7 thrust earthquake occurred along the southeast Loyalty Islands [45], 217 general earthquake notifications about the event were published on Twitter. This earthquake sequence then led to 25 tweets that claimed to predict the next megathrust earthquake in or near this area.

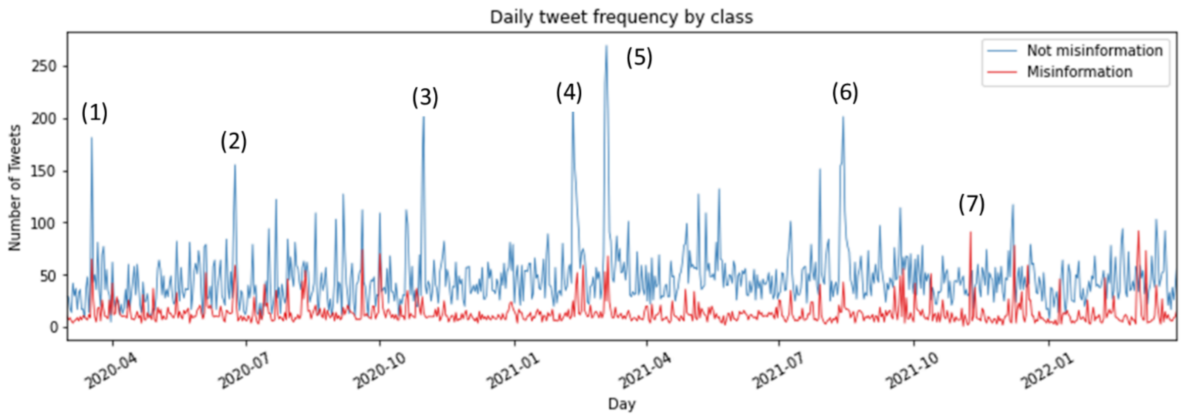


Figure 11. Daily frequency of tweets for earthquake discussion.

Looking at the usage of URLs and media (e.g., pictures and videos) in tweets is relevant to determine how each group reinforces its stance in the debate on Twitter. We processed and analyzed the usage of URLs and media in the tweets

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

of both groups. URLs were reduced to their base domain names and combined the aliases of some major websites (e.g., nytim.es and nytimes.com). Domains that cannot provide any value to the analysis were ignored (e.g., twitter.com, URL-shortening services, and systems for content management).

Investing the number of tweets in each group that made use of media found that within the not-misinformation as well as the misinformation group, about 20% to 25% of the tweets contained media, thus most of the tweets did not contain any media element and no significant differences was found between the groups regarding the media in the tweets.

In case of the URL, Figure 12 presents the number of tweets that linked at least one URL for both the not-misinformation and misinformation groups. 70.7% of the misinformation tweets contained one or more URLs. The four most mentioned URLs were: emsc-csem.org , seismo.info, quakeprediction.com, and youtube.com. The second and third URL domains are privately-run websites that claim to predict or forecast earthquakes.

In comparison, 85.3% of the not-misinformation tweets contained one or more URLs (Figure 12). The most included URLs are recognized and reliable reporting websites and official authorities, i.e., usa.gov, gdacs.org, emsc-csem.org. The second most linked domain is google.com.

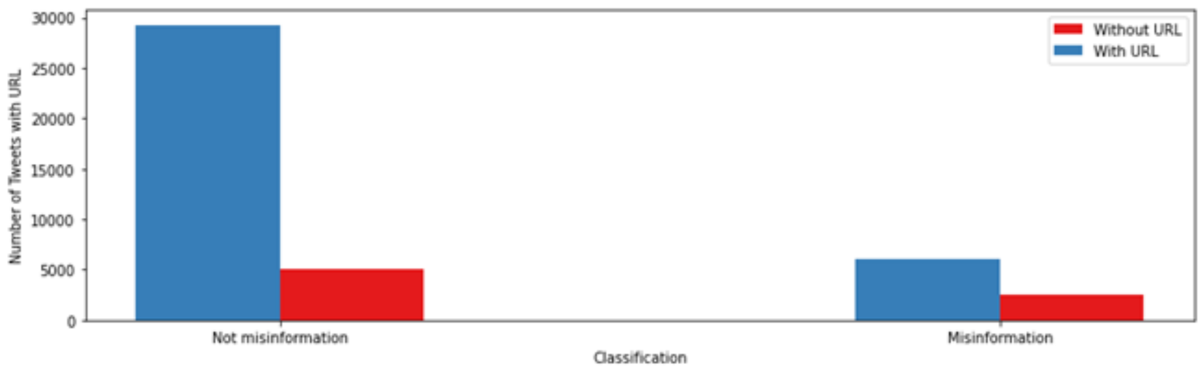


Figure 12. Usage of URLs in the not-misinformation and misinformation tweets.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

3.4. Disaster information seeking behaviour: analysis of disaster pages on Wikipedia

Analysis of Wikipedia's traffic data has proven to be productive for research and has been used in various research areas. Previous research used Wikipedia traffic for building a model for electoral prediction [46]. In the case of disasters and manmade incidents Kanhabua, Nguyen [47] analyzed long-term dynamics of Wikipedia as a global memory place for high-impact events.

Some research concentrated on tools and software dedicated to providing better accessibility to the data and assessing its reliability. Roy, Bhatia [48] investigated information asymmetry in Wikipedia by introducing WikiCompare, a browser plugin that aims at the differences between pages on the same topic in different languages by providing readers with a comprehensive overview of topics by incorporating missing information from Wikipedia pages in other languages. Vardi, Muchnik [49] introduced WikiShark, an online tool that analyzes Wikipedia traffic and trends by extracting the data from the Wikipedia API. Tracking searches on Wikipedia can serve as indication for public's interests and risk perceptions, as well as preferences over time. In this section we analyzed the page traffic of seven representative emergency and disasters case studies (see Table 2, section 2.2)

Data Analysis: We collected the traffic data for the seven representative case studies in multiple languages. For each case study, we used traffic data of the case study's page in its official and usually English as well as a global language (see Table 2, section 2.2). The data can be sliced and displayed in infinite possible ways. This section shall provide some examples for peaks in interest and possible explanations for their cause interest.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

`

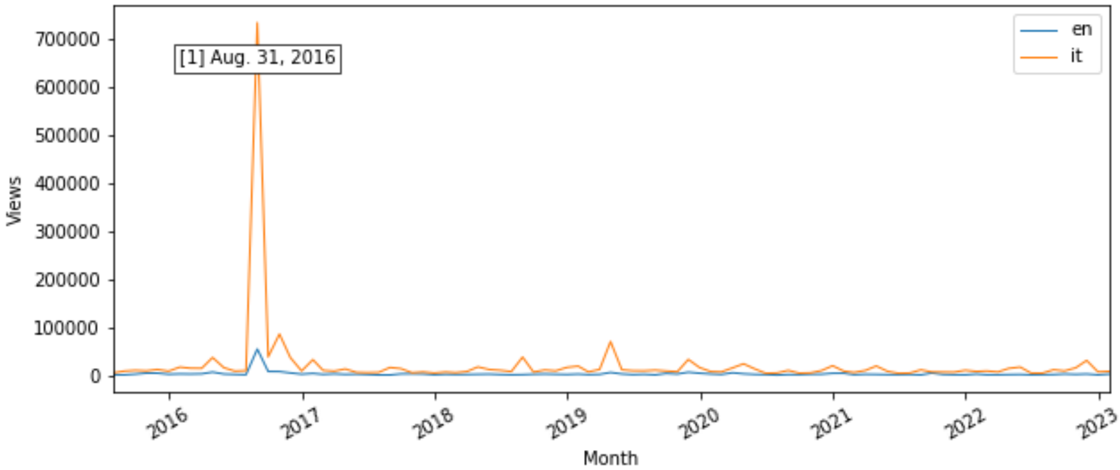


Figure 13. Monthly page views frequency for the L'Aquila earthquake page

The L'Aquila earthquake took place on April 6, 2009 in Italy, with a magnitude of 5.9M, with thousands of foreshocks and aftershocks. 308 casualties have been reported, making this the deadliest earthquake in the region since 1980. Errore. L'origine riferimento non è stata trovata. displays the daily views' frequency for the L'Aquila earthquake page on Wikipedia in English and Italian. A major peak in interest was observed on August 24, 2016 (Errore. L'origine riferimento non è stata trovata., [1]), following another earthquake in a very close region at a magnitude of 6.2 on the moment magnitude scale.

The Manchester Arena bombing is a terror attack that took place on May 22, 2017 in Manchester, United Kingdom, carried by an Islamist suicide bomber. 23 casualties and over a thousand injuries have been reported, with hundreds more suffering from psychological trauma, making it the deadliest terror attack in the UK. **Errore. L'origine riferimento non è stata trovata.** displays the monthly views frequency for the Manchester Arena bombing page on Wikipedia in English. A recurring, seasonal peak in the frequency can be observed in May of every year since 2017 (**Errore. L'origine riferimento non è stata trovata.**, [1-5]), the month the attack took place. The largest peak of May 2018 (Errore. L'origine riferimento non è stata trovata., [1]), may be attributed to increased media coverage of the event after one year of its occurrence, with a series of media reports with updates about the victims and claims of misconduct of journalists when approaching victims' families [50].

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

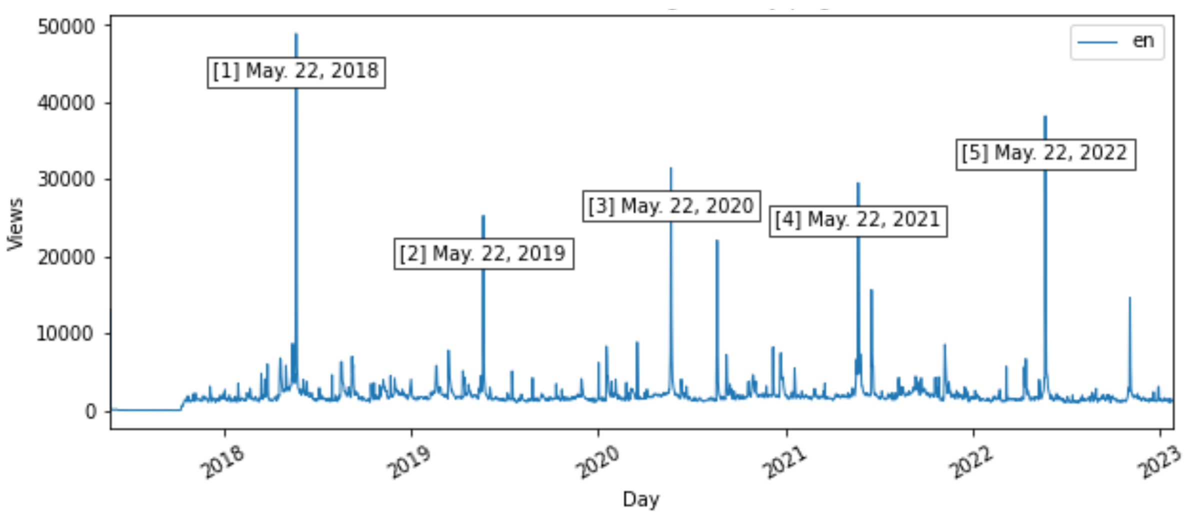


Figure 14. Daily views frequency for the Manchester Arena bombing page in English

The Aude River flooding took place in 2018 in France, caused by heavy thunderstorms leading to 7 meters rise in the height of the river. This was reported to have been the highest level of the river since 1891. At least 14 casualties have been reported. The peak of October 22, 2019 (**Errore. L'origine riferimento non è stata trovata.**, [1]) is likely related to some 120 mm of rain falling in less than 3 hours, causing a river that runs through a city in southern France to rise by almost 4.3 meters by the next day, to a dangerous level [51].

January 22, 2020 has also seen a substantial peak (**Errore. L'origine riferimento non è stata trovata.**, [2]), attributed to Storm Gloria that hit Spain and southern France between January 17, 2020 and January 25, 2020, causing multiple fatalities and damage. Forecasters have said the storm was the worst to hit the region during the winter period since 1982 [52]. The peak of September 8, 2021 (**Errore. L'origine riferimento non è stata trovata.**, [3]), is likely related to nearly two months' worth of rain that fell within just a few hours in southwestern France [53].

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

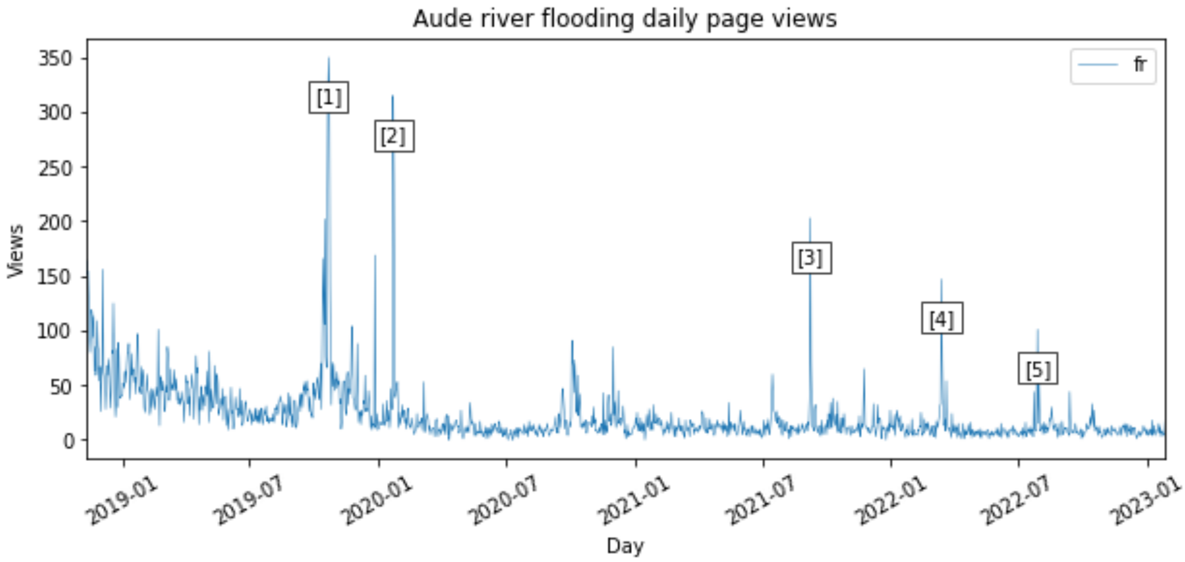


Figure 15. Daily views frequency for the Aude River flooding page in French.

The Visakhapatnam gas leak took place on May 7, 2020 in India, because of an industrial accident at a chemicals plant. 13 casualties have been reported, and over a thousand more reported to be sick after being exposed to the gas. The Visakhapatnam gas leak page in Hindi has seen steady level of interest with an average daily frequency of slightly over 27 views per day since the leak occurred. **Errore. L'origine riferimento non è stata trovata.** displays the daily views' frequency for the Visakhapatnam gas leak page in Hindi. The peak observed on December 23, 2020 (**Errore. L'origine riferimento non è stata trovata.**, [1]), with over 2,200 views in a single day, may be attributed to a second leak of Ammonia gas at the same place on that date. Media reports about the second leak described public concern and fear that it would be as harmful as the first gas leak [54]. The same peak is observed at the daily views' time series of the English page for the incident (**Errore. L'origine riferimento non è stata trovata.**).

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

`

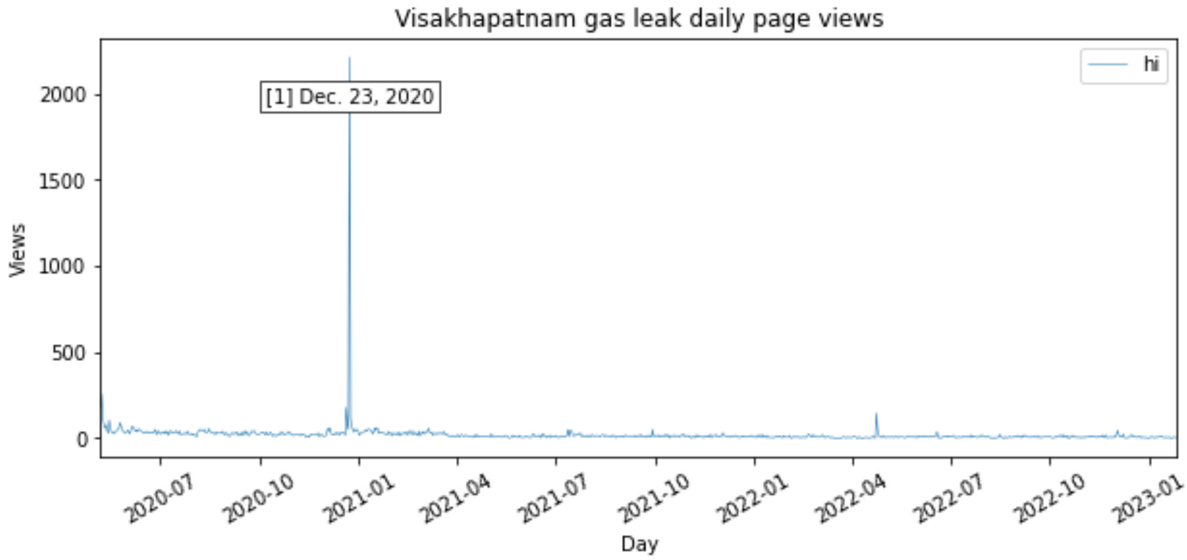
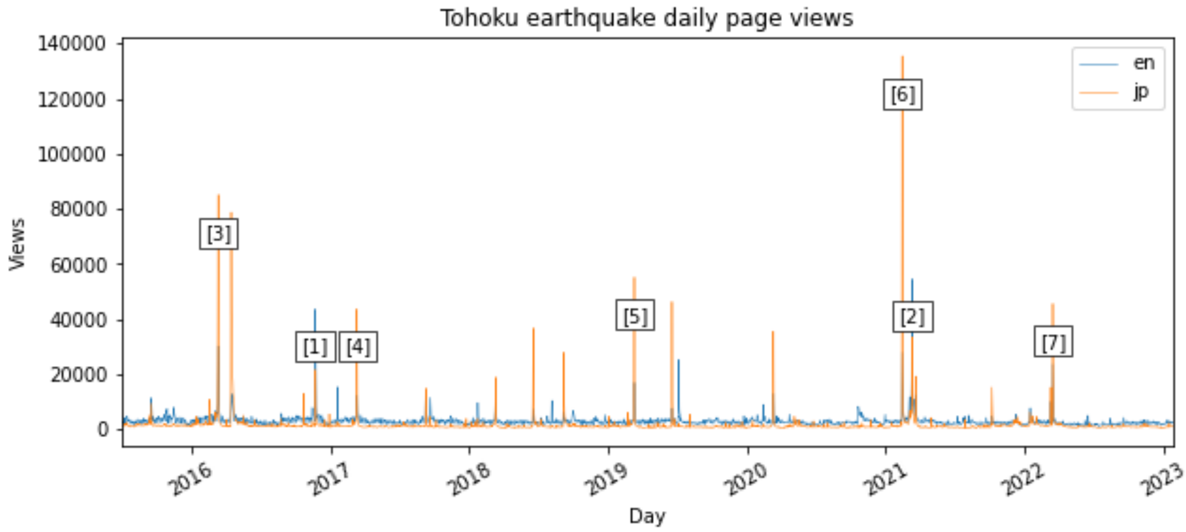


Figure 16. Daily views frequency for the Visakhapatnam gas leak page in Hindi

The Tohoku earthquake and tsunami took place in Japan on March 11, 2011, with a magnitude of 9.1. A tsunami followed the six-minutes earthquake, along with over a dozen thousand aftershocks. Nearly 20 thousand casualties have been, over 6 thousand injuries and more than 2,500 missing have been reported. **Errore. L'origine riferimento non è stata trovata.** displays the daily views' frequency of the Tohoku earthquake and tsunami page in Japanese. Peaks in interest are seen every year on the same day as the actual event, March 11 (**Errore. L'origine riferimento non è stata trovata.**, [2-5]) Other peaks are likely to be associated with subsequent similar events. A couple of peaks in November 21 and November 23, 2016 (**Errore. L'origine riferimento non è stata trovata.**, [1]), may be related to another earthquake near Fukushima, Japan, with a magnitude of 6.9 [55]. The more significant peak of February 13, 2021 (**Errore. L'origine riferimento non è stata trovata.**, [6]), is likely related to a magnitude 7.1 earthquake near Fukushima, Japan, on that day [56]. The more recent peak of March 16, 2022 (**Errore. L'origine riferimento non è stata trovata.**, [7]), can be attributed to a magnitude 7.4 earthquake near Fukushima, Japan.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

`



.

Figure 17. Daily views frequency for the Tohoku earthquake page. Annotations [1-2] are English; [3-7] are Japanese

The COVID-19 pandemic is an ongoing pandemic started in caused by the SARS-CoV-2 virus that has been spreading globally. The pandemic is reported to have caused over 360 million confirmed cases. More than 5.6 million deaths have been reported, but the real amount is estimated to be between 13.4 and 22.7 million deaths. **Errore. L'origine riferimento non è stata trovata.** displays the daily views' frequency of the COVID-19 pandemic English page.

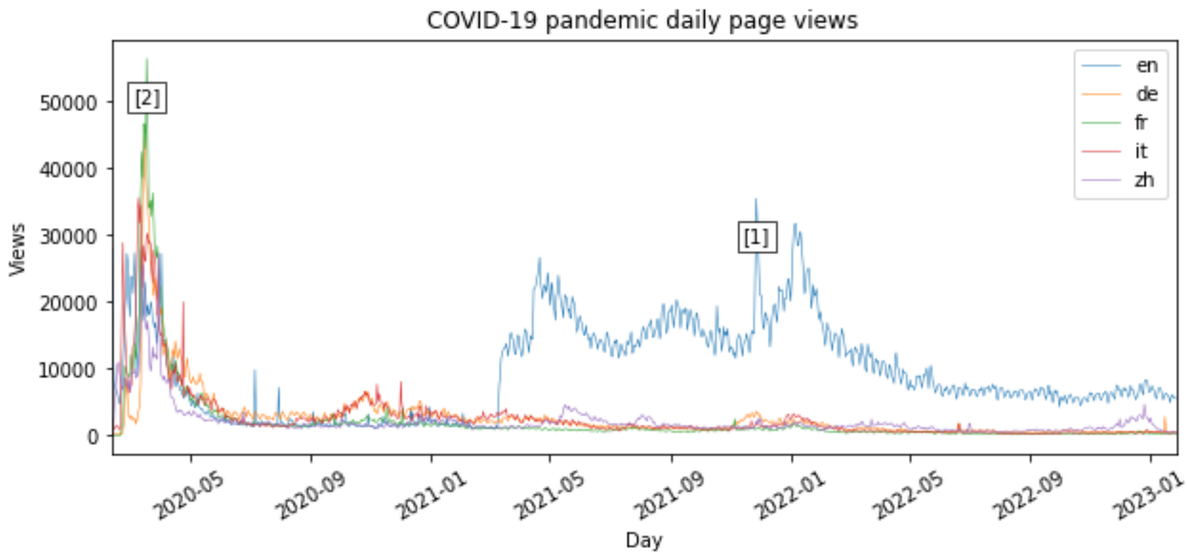


Figure 18. Daily views frequency for the COVID-19 pandemic page in all languages

The peak in views of the English page was observed on November 27, 2021 (**Errore. L'origine riferimento non è stata trovata.**, [1]), with 35,472 page views, roughly double the number of views before and after that date, is likely

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

related to a press release of the World Health Organization on November 28, 2021, stating that the Omicron variant was designated on November 26, 2021, with inherent uncertainty of the disease [57]. The low number of views of the COVID-19 page in Chinese considering the relative size of the population who speak Chinese, and the outbreak taking place in China, is likely related to Wikipedia being inaccessible in China.

The 2021 European floods occurred in July 2021, affecting several European countries throughout the continent. As a result, 242 casualties have been reported, and damage caused by the floods are estimated at a minimum of €10 million. **Errore. L'origine riferimento non è stata trovata.** displays the daily page views for the European floods pages in English and German.

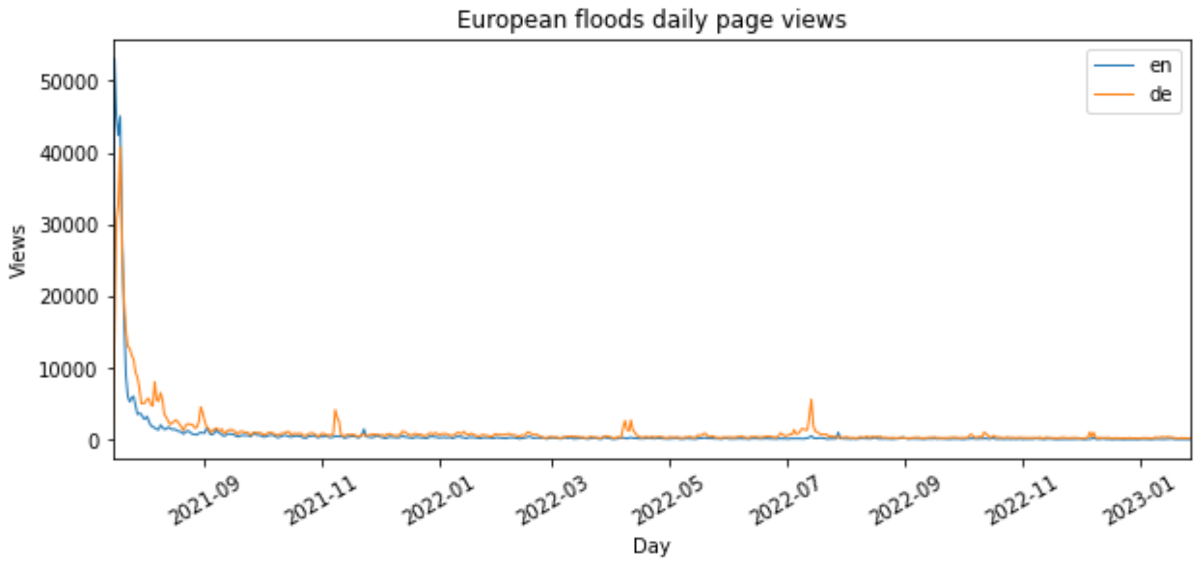


Figure 19. Daily views frequency for the European floods' pages in all languages

We tested the cross-correlation between the views of pages discussing the same event in different languages at different lags. The English page for the COVID-19 virus was not found to be strongly correlated with any of the other pages in other languages. However, we found a very strong cross-correlation between the pages in German, French, and Italian, all of which are languages spoken mainly in bordering countries in Europe. The German page shares a correlation coefficient of r=0.941 at lag *t+1* with the French page, and r=0.913 at lag *t+5* with the Italian page. This finding suggests that page views frequency of the German page is reflected on the French page at the next day, and on the Italian page after 5 days. A strong cross-correlation of r=0.933 was also found between the French and Italian pages at lag *t+3*, and r=0.857 at lag *t-2*, such that the cross-correlation is bidirectional. This finding supports

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

the conclusion that an increase in the page views frequency of either page affects the page views frequency of the other page after a couple of days.
The page in Chinese is also cross-correlated with the European languages. The cross-correlation coefficients between the Chinese page and the German, French, and Italian pages are r=0.816 at lag $t+3$ and r=0.878 at lag $t+2$, and r=0.835 at lag $t+6$, respectively. Meaning that the page views frequency for these European languages in the next few days are following the page views of the Chinese page today.
A possible explanation is that geographically distanced countries that speak the same language, e.g., the United Kingdom and Australia, were affected by the pandemic at different times, thus affecting the reliability of the results. Another explanation is that the information flow for the pandemic starts with English, and other languages come later. Future works may attempt to determine the cause for the correlation between COVID-19 pages in different languages and the lack of correlation between others.
The cross correlation between the EU floods page in English and German was found to be significant and strong with an r=0.932, at lags $t-1$. This result suggests that a higher views frequency on the German page at time t leads to a higher views frequency at the following days on the English page, meaning that the English page views frequency follows the German page the next day.
The correlation coefficient between the L'Aquila earthquake pages in Italian and in English was found to be strong and significant (r=0.961).
A high correlation of r=0.61 (with no significant cross-correlation) was found between the Tohoku earthquake and tsunami pages in English and Japanese. Similar behaviour is seen at the English and Hindi pages for the Visakhapatnam gas leak pages, with a significant positive correlation of r=0.234, but no significant cross correlation.
These results show that people respond to global disasters even if not in their country.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

4. CONCLUSIONS

In this deliverable, we investigated the communication patterns during and after disaster in social media by extracting data from Twitter and Wikipedia.

We analyzed the COVID-19 conspiracy theories discussion on Twitter. The results of the analysis have helped us identify patterns and categorize existing conspiracy theories into four groups. The first – 5G and FilmYourHospital – played a major role at the beginning of the pandemic and then declined sharply. The second – vaccines and Big Pharma – began to play a major role later as vaccines began to be actively introduced. The third – exaggeration and the role of Bill Gates – remained relatively high over a long period of time with some fluctuations. The fourth – GMO and biological weapon – had two peaks and were driven by two events – the emergence of the pandemic and the active start of the vaccination campaign. This shows that many people react to new, unexpected, and incomprehensible risks by resorting to conspiracy narratives. According to the heuristic-systematic model, individuals act heuristically in response to threats, that is, based on their emotions, and according to the group epistemological theory, these individuals also group together around similar labels. This is perfectly visible in the emergence of COVID-19 conspiracy theories. However, when the picture becomes clearer and more reliable and clear information emerges, and many are confronted with the virus themselves, most narratives fade away on their own. But there are some that remain quite persistent.

In addition, we analyzed the 5G conspiracy theory across Europe countries to conclude that the discussion is following the same patterns and is highly correlated between the different languages with different lags. Further analysis on the COVID-19 5G conspiracy theory on Twitter presents a workflow to collect, classify, and analyze tweets related to COVID-19 conspiracy theories using NLP and machine learning methods. The classification process uses two sets of features: sentence embeddings using CT-BERT and SBERT, and external features, and classifies each tweet as supporting the conspiracy, opposing the conspiracy, or neutral/irrelevant. The analysis of the classified dataset raised interesting conclusions. Most basically, there are more tweets opposing the conspiracy theory than there are tweets supporting it. The conspiracy theory appears to be declining, and while supporters of the conspiracy theory initially dominated the

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

conversation during the first months, opponents took over the conversation shortly after

For the Monkeypox dataset we found that only a third of the tweets related to monkeypox discuss misinformation, whereas two thirds of the tweets are neutral. Analysis of the users that participated in the discussion revealed that the interest is driven by users that participate in the discussion for the first time. We also found that tweets countering misinformation referenced authoritative sources such as the CDC and WHO more often. As such, it may be recommended to provide more frequent updates on authoritative websites with reliable information for use by the users countering misinformation and indirectly support the fight in misinformation. On the other hand, tweets that spread misinformation referenced mostly websites that host user content, such as YouTube. Encouraging these platforms to monitor content and adopt stricter community guidelines, would likely reduce the amount of misinformation to be shared. Additionally, it may be recommended to notifying users of social media platforms about the sources of the content and their credibility, to minimize unaware echoing of misinformation.

The earthquakes predictions analysis identified that unlike an event-driven misinformation or conspiracy, such as those related to COVID-19 where the discussion dissipates over time, earthquakes carry inherent uncertainty as they occur randomly, cannot be mitigated, and cannot be predicted, thus reigniting the discussion time after time. However, we also show that after several events, earthquake predications claims are more often spread on Twitter, which is problematic since affected people try to give sense to the on-going earthquake sequence [58]. Especially after strong events, institutions responsible for the public communication need to provide rapid, accurate information about what is on-going so that people do not fill the information void with false information. Regarding the elements to support one's statements, we identified that people on Twitter rather use URLs and not media elements to underline their arguments. In comparison to the 5G analysis who showed that COVID-19 supporter's tweets of the conspiracy theory contain more URLs than tweets from the opponents, we did not see this pattern for earthquake predictions. This might be explained by the fact that for the 5G conspiracy and COVID-19 not much accurate information material to share was available, whereas for earthquakes on national Seismological Services' websites scientific information clarifying that predictions are not possible are available.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

`

Wikipedia typically hosts pages relevant to multiple communities in multiple languages. With the assumption that users are more likely to view pages in their native language and depending on their physical location, we collected traffic data of seven representative disasters and emergency case studies on Wikipedia. We analyzed the data to find abnormalities, such as peaks in interest. Then, we provided possible explanations that could assist future researchers in better understanding the unfolding of the public interest in emergencies and disasters.

In general, the results all shed light on communication patterns after emergencies and disasters. People tend to adhere to conspiracy theories and misinformation to fill in a vacuum of Information, especially during periods of uncertainty. The results show that this tendency is common across various countries and cultures. Further, combining the results of the Twitter analysis with the Wikipedia analysis reveals that the need for information is common in various languages and countries. People seek for information after significant events to learn about its impact and consequences. These patterns are important for understanding what leads people to certain information sources, what they consider reliable information, and how they consume information during and after emergencies and disasters.

The results presented in this research allow us to make recommendations to policy makers – especially in the area of what information is consumed by the public, what are the information needs, and what is the preferred channel to provide it. As we have shown, information provided by official institutes and authorities is useful for "fighting" against misinformation. Enhancing the ability of such institutions to provide accurate and timely information is important for filling the information gaps, especially in the first moments. However, building such capacity requires investments in the administrative (in terms of staffing) and technology (in terms of detecting misinformation), which should be developed before the disaster. Further recommendations will follow soon in the next deliverable. Also, further research is needed to understand the effect of using URL links and Media in tweets on risk perception.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

`

REFERENCES

1. Bier, V.M., *On the state of the art: risk communication to the public.* Reliability Engineering & System Safety, 2001. **71**(2): p. 139-150.
2. Coyle, D. and P. Meier, *New technologies in emergencies and conflicts: The role of information and social networks*, in *New Technologies in Emergencies and Conflicts: The Role of Information and Social Networks*. 2009. p. 54-54.
3. Norris, F.H., et al., *Community Resilience as a Metaphor, Theory, Set of Capacities, and Strategy for Disaster Readiness.* American Journal of Community Psychology, 2008. **41**(1): p. 127-150.
4. Lindell, M.K. and D.J. Whitney, *Correlates of Household Seismic Hazard Adjustment Adoption.* Risk Analysis, 2000. **20**(1): p. 13-26.
5. Joffe, H., et al., *Social Representations of Earthquakes: A Study of People Living in Three Highly Seismic Areas.* Earthquake Spectra, 2013. **29**(2): p. 367-397.
6. Solberg, C., T. Rossetto, and H. Joffe, *The social psychology of seismic hazard adjustment: re-evaluating the international literature.* Nat. Hazards Earth Syst. Sci., 2010. **10**(8): p. 1663-1677.
7. Paton, D., et al., *Predicting community earthquake preparedness: a cross-cultural comparison of Japan and New Zealand.* Natural Hazards, 2010. **54**(3): p. 765-781.
8. Mileti, D.S., et al., *Annotated bibliography for public risk communication on warnings for public protective actions response and public education*(revision 4).* Natural Hazards Centre, University of Colorado at Boulder, 2006.
9. Kirschenbaum, A., C. Rapaport, and D. Canetti, *The impact of information sources on earthquake preparedness.* International Journal of Disaster Risk Reduction, 2017. **21**: p. 99-109.
10. Lindell, M.K. and R.W. Perry, *The Protective Action Decision Model: Theoretical Modifications and Additional Evidence.* Risk Analysis, 2012. **32**(4): p. 616-632.
11. Verrucci, E., et al., *Digital engagement methods for earthquake and fire preparedness: a review.* Natural Hazards, 2016. **83**(3): p. 1583-1604.
12. Wood, M.M., et al., *Communicating Actionable Risk for Terrorism and Other Hazards⋆.* Risk Analysis, 2012. **32**(4): p. 601-615.
13. Kyne, D. and W. Donner, *Kyne–Donner model of authority's recommendation and hurricane evacuation decisions: A study of hypothetical hurricane event in the Rio Grande Valley, Texas.* Population Research and Policy Review, 2018. **37**(6): p. 897-922.
14. Zhu, W., J. Wei, and D. Zhao, *Anti-nuclear behavioral intentions: The role of perceived knowledge, information processing, and risk perception.* Energy Policy, 2016. **88**: p. 168-177.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

15. Hobbs, J., et al., *Communicating Health Information to an Alarmed Public Facing a Threat Such as a Bioterrorist Attack.* Journal of Health Communication, 2004. **9**(1): p. 67-75.
16. Schmalzried, H.D., L. Fleming Fallon Jr., and E.A. Harper, *Assessing informational website communications during emergencies and disasters.* International Journal of Nonprofit and Voluntary Sector Marketing, 2012. **17**(3): p. 199-207.
17. Simon, T., A. Goldberg, and B. Adini, *Socializing in emergencies—A review of the use of social media in emergency situations.* International Journal of Information Management, 2015. **35**(5): p. 609-619.
18. Cassa, C.A., et al., *Twitter as a sentinel in emergency situations: lessons from the Boston marathon explosions.* PLoS Curr, 2013. **5**.
19. Taylor, M., et al., *The role of social media as psychological first aid as a support to community resilience building.* Australian Journal of Emergency Management, The, 2012. **27**(1): p. 20-26.
20. Kavanaugh, A., et al., *Social media use by government: from the routine to the critical*, in *Proceedings of the 12th Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times*. 2011, Association for Computing Machinery: College Park, Maryland, USA. p. 121–130.
21. Huang, C.-M., E. Chan, and A.A. Hyder, *Web 2.0 and Internet Social Networking: A New tool for Disaster Management? - Lessons from Taiwan.* BMC Medical Informatics and Decision Making, 2010. **10**(1): p. 57.
22. Merchant, R.M., S. Elmer, and N. Lurie, *Integrating Social Media into Emergency-Preparedness Efforts.* New England Journal of Medicine, 2011. **365**(4): p. 289-291.
23. Erokhin, D., A. Yosipof, and N. Komendantova, *COVID-19 Conspiracy Theories Discussion on Twitter.* Social Media + Society, 2022. **8**(4): p. 20563051221126051.
24. Elroy, O. and A. Yosipof. *Analysis of COVID-19 5G Conspiracy Theory Tweets Using SentenceBERT Embedding*. in *Artificial Neural Networks and Machine Learning – ICANN 2022*. 2022. Bristol UK: Springer Nature Switzerland.
25. Piskorski, J., J. Haneczok, and G. Jacquet. *New Benchmark Corpus and Models for Fine-grained Event Classification: To BERT or not to BERT?* in *Proceedings of the 28th International Conference on Computational Linguistics*. 2020.
26. González-Carvajal, S. and E.C. Garrido-Merchán, *Comparing BERT against traditional machine learning text classification.* arXiv preprint arXiv:2005.13012, 2020.
27. Micallef, N., et al. *The role of the crowd in countering misinformation: A case study of the COVID-19 infodemic*. in *2020 IEEE International Conference on Big Data (Big Data)*. 2020. IEEE.
28. Beskow, D.M. and K.M. Carley. *Bot-hunter: a tiered approach to detecting & characterizing automated activity on twitter*. in *Conference paper. SBP-*

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

BRiMS: International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation. 2018.
29. ODonovan, J., et al. *Credibility in context: An analysis of feature distributions in twitter*. in *2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Confernece on Social Computing*. 2012. IEEE.
30. Gupta, A., et al. *Tweetcred: Real-time credibility assessment of content on twitter*. in *International conference on social informatics*. 2014. Springer.
31. Devlin, J., et al., *Bert: Pre-training of deep bidirectional transformers for language understanding.* arXiv preprint arXiv:1810.04805, 2018.
32. Kovaleva, O., et al., *Revealing the dark secrets of BERT.* arXiv preprint arXiv:1908.08593, 2019.
33. Müller, M., M. Salathé, and P.E. Kummervold, *Covid-twitter-bert: A natural language processing model to analyse covid-19 content on twitter.* arXiv preprint arXiv:2005.07503, 2020.
34. Reimers, N. and I. Gurevych, *Sentence-bert: Sentence embeddings using siamese bert-networks.* arXiv preprint arXiv:1908.10084, 2019.
35. Hutto, C. and E. Gilbert. *Vader: A parsimonious rule-based model for sentiment analysis of social media text*. in *Proceedings of the international AAAI conference on web and social media*. 2014.
36. Liu, Y., et al., *Roberta: A robustly optimized bert pretraining approach.* arXiv preprint arXiv:1907.11692, 2019.
37. Adoma, A.F., N.-M. Henry, and W. Chen. *Comparative analyses of bert, roberta, distilbert, and xlnet for text-based emotion recognition*. in *2020 17th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*. 2020. IEEE.
38. Tarunesh, I., S. Aditya, and M. Choudhury, *Trusting roberta over bert: Insights from checklisting the natural language inference task.* arXiv preprint arXiv:2107.07229, 2021.
39. Naseer, M., M. Asvial, and R.F. Sari. *An empirical comparison of bert, roberta, and electra for fact verification*. in *2021 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*. 2021. IEEE.
40. Nguyen, D.Q., T. Vu, and A.T. Nguyen, *BERTweet: A pre-trained language model for English Tweets.* arXiv preprint arXiv:2005.10200, 2020.
41. World Health Organization. *Questions and Answers: Monkeypox*. 2022 February 2, 2023]; Available from: https://www.who.int/news-room/questions-and-answers/item/monkeypox.
42. World Health Organization. *Monkeypox - United Kingdom of Great Britain and Northern Ireland* 2022 February 2, 2023]; Available from: https://www.who.int/emergencies/disease-outbreak-news/item/2022-DON381.
43. Main, I., *Is the reliable prediction of individual earthquakes a realistic scientific goal.* Nature, 1999. **397**(1).

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

44. Fallou, L., et al., *How to Fight Earthquake Misinformation: A Communication Guide*. 2022. p. 2418-2422.
45. Ye, L., et al., *Shallow Megathrust Rupture during the 10 February 2021 M w 7.7 Southeast Loyalty Islands Earthquake Sequence*. The Seismic Record, 2021. **1**(3): p. 154-163.
46. Yasseri, T. and J. Bright, *Wikipedia traffic data and electoral prediction: towards theoretically informed models*. EPJ Data Science, 2016. **5**(1): p. 1-15.
47. Kanhabua, N., T.N. Nguyen, and C. Niederée. *What triggers human remembering of events? A large-scale analysis of catalysts for collective memory in Wikipedia*. in *IEEE/ACM Joint Conference on Digital Libraries*. 2014. IEEE.
48. Roy, D., S. Bhatia, and P. Jain, *Information asymmetry in Wikipedia across different languages: A statistical analysis*. Journal of the Association for Information Science and Technology, 2022. **73**(3): p. 347-361.
49. Vardi, E., et al. *WikiShark: An Online Tool for Analyzing Wikipedia Traffic and Trends*. in *Companion Proceedings of the Web Conference 2021*. 2021.
50. BBC News. *Manchester Arena attack: Bomb 'injured more than 800'*. 2018 January 24, 2022]; Available from: https://www.bbc.com/news/uk-england-manchester-44129386.
51. Floodist. *France – Major Flooding in South After 120mm of Rain in 3 Hours*. 2019 January 30, 2022]; Available from: https://floodlist.com/europe/france-south-flood-herault-october-2019.
52. The Watchers. *Storm Gloria: Death toll rises to 13, over 1 500 evacuated in France after two months' worth of rain in 48 hours*. 2020 January 30, 2022]; Available from: https://watchers.news/2020/01/24/storm-gloria-spain-france-january-2020/.
53. Floodist. *France – Record Rainfall Floods Streets in Lot-Et-Garonne*. 2021; Available from: https://floodlist.com/europe/france-floods-agen-lotetgaronne-september-2021.
54. India TV News. *Andhra Pradesh: Ammonia gas leak at private company in Visakhapatnam*. January 27, 2022]; Available from: https://www.indiatvnews.com/news/india/andhra-pradesh-ammonia-gas-leak-at-visakhapatnam-private-company-706636.
55. United States Geological Survey. *Manitude 6.9 Earthquake off Japan*. 2016 January 30, 2022]; Available from: https://www.usgs.gov/news/featured-story/magnitude-69-earthquake-japan.
56. Wikipedia. *2021 Fukushima earthquake*. 2022 January 30, 2022]; Available from: https://en.wikipedia.org/wiki/2021_Fukushima_earthquake.
57. World Health Organization. *Update on Omicron*. 2021 January 24, 2022]; Available from: https://www.who.int/news/item/28-11-2021-update-on-omicronRetrieved.
58. Jahanbakhsh, F., et al., *Exploring lightweight interventions at posting time to reduce the sharing of misinformation on social media*. Proceedings of the ACM on Human-Computer Interaction, 2021. **5**(CSCW1): p. 1-42.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

Report about communication patters

D 7.1

`

sCience and human factOr for Resilient sociEty

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 101021746. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.