Monitoring user opinions and side effects on COVID-19 vaccines in the Twittersphere: Infodemiology Study of Tweets

Beatrice Portelli, Simone Scaboro, Roberto Tonino, Emmanuele Chersoni, Enrico Santus, Giuseppe Serra

Submitted to: Journal of Medical Internet Research
on: November 22, 2021

Table of Contents

Original Manuscript ... 5
Supplementary Files ... 29
Figures ... 30
Figure 1 ... 31
Figure 2 ... 32
Figure 3 ... 33
Figure 4 ... 34
Figure 5 ... 35
Figure 6 ... 36
Figure 7 ... 37
Figure 8 ... 38
Figure 9 ... 39
Figure 10 ... 40
Figure 11 ... 41
Figure 12 ... 42
Figure 13 ... 43
TOC/Feature image for homepages ... 44
TOC/Feature image for homepage 0 ... 45

Monitoring user opinions and side effects on COVID-19 vaccines in the Twittersphere: Infodemiology Study of Tweets

Beatrice Portelli[1] MSc; Simone Scaboro[1] MSc; Roberto Tonino[1] BSc; Emmanuele Chersoni[2]; Enrico Santus[3]; Giuseppe Serra[1]

[1]Department of Mathematics, Computer Science and Physics University of Udine Udine IT
[2]The Hong Kong Polytechnic University Hung Hom HK
[3]Decision Science and Advanced Analytics for MAPV & RA, Bayer Bayer Pharmaceuticals Whippany US

Corresponding Author:
Beatrice Portelli MSc
Department of Mathematics, Computer Science and Physics
University of Udine
via delle Scienze 206
Udine
IT

Abstract

Background: In the current phase of the COVID-19 pandemic, we are witnessing the most massive vaccine rollout in human history. Like any other drug, vaccines may cause unexpected side effects, which need to be timely investigated to minimize harm in the population. If not properly dealt with, side effects may also impact the public trust in the vaccination campaigns carried out by the national governments.

Objective: Monitoring social media for the early identification of side effects and understanding the public opinion on the vaccines are of paramount importance to ensure a successful and harmless rollout. The objective is to create a web portal to monitor the opinion of social media users on the vaccines, to provide a tool for journalists, scientists, and users alike to visualize how the general public is reacting to the vaccination campaign.

Methods: In this paper, we present a tool to analyze the public opinion on COVID-19 vaccines from Twitter, exploiting, among the others: a state-of-the-art system for the identification of Adverse Drug Events (ADEs) on social media; Natural Language Processing models for sentiment analysis; statistical tools and open-source databases to visualize the trending hashtags, news articles and their factuality. All the modules of the system are displayed through a web portal available at http://ailab.uniud.it/covid-vaccines/.

Results: A set of 650,000 tweets have been collected and analyzed starting from December 2020. Tweet collection and analysis is an ongoing process. The results of the analysis are made public on a web portal (updated daily), together with a description of the processing methods and ways to access the preprocessed data. The collected data provide sensible insights in the public opinion on the vaccines and how their main worries changed in time. They show how news coverage had a high impact on the set of topics discussed by Twitter users, and the reassuring trend that users have a high tendency to only share news from reliable sources when discussing COVID-19 vaccines.

Conclusions: We presented a tool connected with a web portal to monitor and display some key aspects of the public's reaction to COVID-19 vaccines. The system also provides an overview of the opinions of the Twittersphere through graphic representations and represents a tool for the extraction of suspected adverse events from tweets with a Deep Learning model.

(JMIR Preprints 22/11/2021:35115)
DOI: https://doi.org/10.2196/preprints.35115

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?
Please make my preprint PDF available to anyone at any time (recommended).
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.

Only make the preprint title and abstract visible.

✓ **No, I do not wish to publish my submitted manuscript as a preprint.**

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Original Paper

Beatrice Portelli, University of Udine, portelli.beatrice@spes.uniud.it
Simone Scaboro, University of Udine, scaboro.simone@spes.uniud.it
Roberto Tonino, University of Udine, tonino.roberto@spes.uniud.it
Emmanuele Chersoni, The Hong Kong Polytechnic University, emmanuele.chersoni@polyu.edu.hk
Enrico Santus, DSIG - Bayer Pharmaceuticals, esantus@gmail.com
Giuseppe Serra, University of Udine, giuseppe.serra@uniud.it

Monitoring user opinions and side effects on COVID-19 vaccines in the Twittersphere: Infodemiology Study of Tweets

Abstract

Background: In the current phase of the COVID-19 pandemic, we are witnessing the most massive vaccine rollout in human history. Like any other drug, vaccines may cause unexpected side effects, which need to be timely investigated to minimize harm in the population. If not properly dealt with, side effects may also impact the public trust in the vaccination campaigns carried out by the national governments.
Objective: Monitoring social media for the early identification of side effects and understanding the public opinion on the vaccines are of paramount importance to ensure a successful and harmless rollout. The objective is to create a web portal to monitor the opinion of social media users on the vaccines, to provide a tool for journalists, scientists, and users alike to visualize how the general public is reacting to the vaccination campaign.
Methods: In this paper, we present a tool to analyze the public opinion on COVID-19 vaccines from Twitter, exploiting, among the others: a state-of-the-art system for the identification of Adverse Drug Events (ADEs) on social media; Natural Language Processing models for sentiment analysis; statistical tools and open-source databases to visualize the trending hashtags, news articles and their factuality. All the modules of the system are displayed through a web portal available at http://ailab.uniud.it/covid-vaccines/.
Results: A set of 650,000 tweets was collected and analyzed in an ongoing process started in December 2020. The results of the analysis are made public on a web portal (updated daily), together with the processing tools and data. The data provide insights on the public opinion about the vaccines and its change in time. For example, users show a high tendency to only share news from reliable sources when discussing COVID-19 vaccines (98% of the shared URLs). The general sentiment of the users towards the vaccines is negative/neutral, but the system is able to record fluctuations in the attitude towards specific vaccines in correspondence with specific events (eg, news about new outbreaks). The data also show how news coverage had a high impact on the set of discussed topics. To further investigate this point, we perform a more in-depth analysis of the data regarding AstraZeneca. We observe how media coverage of blood-clot related side effects suddenly shifted the topic of public discussions regarding both AstraZeneca and the other vaccines. This is particularly evident when visualizing the most frequently discussed symptoms for the vaccines and comparing them month-by-month.
Conclusions: We presented a tool connected with a web portal to monitor and display some key aspects of the public's reaction to COVID-19 vaccines. The system also provides an overview of the opinions of the Twittersphere through graphic representations and represents a tool for the extraction of suspected adverse events from tweets with a Deep Learning model.

Keywords: Adverse Drug Events; COVID-19; Digital Pharmacovigilance; Opinion Mining;

Vaccines

Introduction

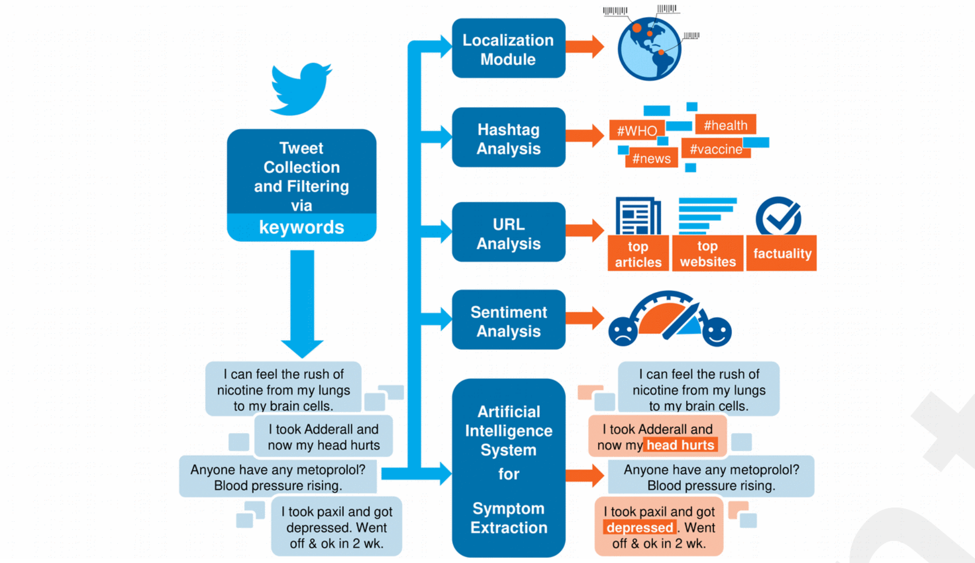
Background

The COVID-19 pandemic has been at the heart of the discussions on all media outlets for almost two years. Those debates touch very important and sensitive topics, such as health, politics, work, school and personal freedom, just to cite a few of them. In a general effort to tackle the pandemic, many countries have engaged in the fastest and most massive vaccine rollout in human history: in less than a year, several vaccines have been created, tested, and distributed around the world, and many others are at the last phase of the clinical trials and/or waiting for the approval from the regulatory agencies [1]. Despite the great efforts put into development, their rollout has been slowed down in various countries [2] due to hesitancy and fake news poisoning the social media debates. The vaccination rollout for the first strains of the virus has proceeded slower than initially planned, but experts agree that it is imperative to find ways to accelerate future iterations to keep pace with the new COVID-19 variants [3]. One of the ways to improve this process is to study how the population reacted to the first vaccination campaigns, which kind of information/misinformation was shared and what impact this had on vaccination hesitancy.

Social media platforms are, of course, one of the main stages of this debate.
In the last years, microblogging services like Twitter have seen an increase in popularity due to their immediacy and ease of use. Also brands, institutional bodies, politicians, public figures and traditional news outlets have realized the importance of having a presence on these platforms, which allow them to deliver messages with high impact and unprecedented reach [4,5].

The rapid spread of the pandemic, the fast development of the vaccines, and the increasing worries about their safety have been hot topics on the social media since the very beginning.
The vaccination campaigns planned by national governments could therefore be seriously hampered by misinformation on such outlets [6,7]. Many recent studies [8] have taken great interest in analyzing different social media to track the sentiment of users about COVID-19 vaccinations across different cities [9], looking for the main misconceptions and complaints about the COVID-19 control measures [10] and the confidence in the efficacy of the vaccines [11].
These are only some examples of why it is highly informative and beneficial to monitor social media platforms to discover health-related issues (eg, detecting mentions of adverse events) and to better understand the public opinion (eg, monitoring the information quality and contrasting the spread of fake news). From this point of view, modern systems for Digital Pharmacovigilance can deploy natural language processing techniques to collect and analyze the online discussions. This allows them to identify potential adverse events (AEs) that may not have been detected during clinical trials and allow timely decisions to reduce their harm. In the near future, it is likely that even public healthcare systems will increase their monitoring activities on social media platforms, with the goal of identifying and treating health issues such as mental diseases, managing information by contrasting fake news or launching prevention campaigns (eg, to mitigate vaccine hesitancy) [12].

Figure 1. Schema of the full system architecture used to analyze the information displayed on the web portal.



Objective

In this contribution, we present an overview of our system for monitoring and analyzing vaccine opinions. Its modules aim at generating insights from Twitter on the topic of COVID-19 vaccines. The tool collects tweets daily and analyzes them to extrapolate information about the public reception of the vaccination campaigns on social media. We also break down the information in easy-to-read charts, for both specialized and general audiences, displaying them on our interactive web portal.

(Figure 1) illustrates the architecture of the full system behind the web portal. It consists of a module dedicated to data collection and various modules dedicated to data processing.
The main features of the system are: (1) Localization; (2) Hashtag Analysis; (3) News Sources Analysis; (4) Sentiment Analysis; (5) Symptom Extraction.

The Symptom Extraction module, in particular, consists of a Deep Learning architecture we created specifically for this task, based on SpanBERT [13], and is one of the state-of-the-art models for Adverse Event Detection [14–16].

Each one of the processing modules is built to extract specific information from the collected tweets (eg, the most used hashtag or the most shared links). This information is then cleaned and provided to the user through the web portal, with interactive charts and diagrams. To ensure greater readability, colors, and shapes were preferred over figures when presenting the data.

To summarize: our objective is to present a tool for the collection and processing of data on COVID-19 vaccines, followed by their visualization on a web dashboard. Differently from other previous works, we focus on monitoring tweets about specific vaccines. This allows us to compare their public reception and how it changed in time. Besides combining various features that can be found separately in recent works, we also introduce innovative modules (eg, symptom extraction), which give us new insights on the public discourse.
The code for the data collection and the preprocessing tools, as well as all the pre-computed statistics and the IDs of the tweets can be openly accessed at https://github.com/AilabUdineGit/covid-vaccines-tools. The amount and kind of data that can be shared openly is limited by the Twitter privacy policy. Further information can be requested

for research purposes. In the Results section we will also present a case study on the AstraZeneca vaccine, as an example of the analyses that can be carried out on the data using our system.

Related Work

Since the start of the COVID-19 pandemic, organizations world-wide have stressed the need to collect and share all data available on the virus, its effects and all related research [17]. As time passed, these resources grew in size, and some researchers started analyzing data coming from social media, too.

For example, in [10] the authors collected 31,100 Australian tweets (from January 20, 2020 to October 22, 2020) related to COVID-19 vaccines. Their paper focuses on analyzing the sentiment and opinion of the users about the vaccines and the main recurring topics in the tweets. Similarly, [9] collected and analyzed reddit comments about COVID-19 vaccines from three Canadian cities (from July 13, 2020 to June 14, 2021) and performed a comparison of the sentiment and main discussion topics among the three locations. Other recent works focused on analyzing sentiment and discussion topics in tweets about COVID-19, generated in other countries and in different time periods [18–20].

These works were carried out on very specific time periods and focused on a single aspect of the social media messages. A more comprehensive work was carried out by the authors of AvaxTweets [21], a public dataset of Twitter posts and accounts that exhibited a strong stance against COVID-19 vaccines, collected between October 2020 and December 2020. The authors analyzed the accounts in terms of the most frequent hashtags, which news sources they shared, and their most likely political orientation, looking for useful insights on how to counter misinformation and vaccine hesitancy.

However both this and the preceding works were carried out on a limited time scale and aimed specifically at the research community, providing no tools or web interfaces to explore the data.

At the same time, various researchers focused not only on data collection, but also on ways to start processing and visualizing the data to make them available for a broader public. COnVIDa [22] is a web-based platform that provides day-to-day interactive information on COVID-19-related conditions in Spain collating data from various sources (e.g., health databases, mortality reports, statistics, information on citizens' mobility from Google and Apple Maps). This project focuses on a single country and tries to combine different aspects of the situation to give the viewer a more complete visualization. CoVaxxy [23] is another dataset and online dashboard that focuses on the correlations between tweets about COVID-19 vaccines, credibility of the shared news and vaccine adoption on U.S. geolocated posts. [24] presents another recent tool that collected and analyzed Twitter conversations from March 1, 2020 to June 5, 2020. The dashboard visualizes sentiment information, trending topics, but focuses especially on the credibility of the news shared in the tweets and on how misinformation spreads.

Our proposed system includes many of the features offered by these separate works, such as: continuous day-to-day data collection and processing (since December 15, 2020); global data collection (not country-specific); sentiment analysis and news sources analysis. Our work differs from the previous ones for the following:

- a focused monitoring of specific vaccines since the date of their approval, which enables the users to compare the crowd's reaction to them;
- a wide variety of processing modules (not focused on a single aspect), to allow a multi-

faced view of the social media discourse;

- a comprehensive dashboard to visualize all the processed data in an easy-to-read way for different categories of users;
- an innovative symptom extraction module to track the most discussed side effects;
- openly available code and data.

Methods

Data Collection

Tweets are collected using the Twitter Application Programming Interface (API) [25]. To recover the most recent tweets mentioning the specific vaccine we use the query “covid vaccine <vaccine_name>”, where <vaccine_name> is the lowercase name of one of the monitored vaccines (originally Pfizer-BioNTech, AstraZeneca, and Moderna, then expanded to include the newly introduced vaccines). We require that all keywords are present in the tweet (either as text, hashtag or as part of a link inside the tweet) and that each query contains the name of only one vaccine.
Tweets are selected among the “most recent”, as opposed to the “most popular”, and retweets are discarded. This is done to avoid skewing the data with popular tweets produced by few influential users. Although we are collecting tweets in various languages, only the ones written in English are passed to the following stages of processing, as most of our current modules are language-dependent. Nonetheless, we are storing these data for future research, as we plan to overcome this limitation in the near future with the introduction of multi-lingual models (in particular for AE detection and sentiment analysis) and automated translation services. This will allow us to perform a complete analysis for all the monitored languages.
The query is run every 24 hours, with a cap of 7,000 requested tweets per day (to be divided among the monitored vaccines) imposed by the limits of the API. Despite the theoretical limitation, the number of new tweets that matched the query in the last 24 hours never exceeded 7,000.

The body of the remaining messages undergoes extra pre-processing steps to identify possible duplicates and discard tweets that are practically identical (apart from hashtags, punctuation, or URLs). This situation happens, for example, when users share a piece of news using the “Share on Twitter” button provided by news websites. If the user simply shares the news without adding any comments (or adding just an hashtag), the result is a great number of nearly-identical tweets. They provide no additional information aside from the fact that the particular piece of news was shared multiple times, so they are marked as “duplicated”. They are not discarded because they can provide useful information on which articles went viral, but they are marked to avoid introducing noise into other kinds of analyses.
De-duplication is performed by removing all hashtags, URLs and punctuation and performing a (fuzzy) matching with the collection of “unique” tweets already collected.

Data collection started on December 10th, 2020, concurrently with the FDA approval of the first COVID-19 vaccine (Pfizer-BioNTech), and the system has currently (September 7th, 2021) analyzed over 650,000 tweets. (Table 1) reports the names of the vaccines tracked at the time of writing and the date we started collecting data about them.

Table 1. Names of the tracked vaccines and dates on which data collection started.

Vaccine name	Start date
Pfizer-BioNTech	10/12/2020
AstraZeneca	11/12/2020
Moderna	16/12/2020
Sinopharm	24/02/2021
Sputnik V	24/02/2021
Sinovac	24/02/2021
Johnson & Johnson	01/04/2021

Tweet storage and handling

Twitter is a major social network and, as such, it has strict policies to regulate the ethical use of its data and the privacy of its users. Following their guidelines, we collect and store only the information needed for the processing steps which are currently implemented. We memorize the outputs of the modules and discard all the sensitive data soon afterward. We also memorize the tweet id, which allows us (and other researchers) to access the original tweet in the future, as long as the user does not delete it or change its visibility.
If a tweet needs to be displayed on a web interface, we use the API provided by Twitter, which allows us to display tweets on demand given their tweet id (and only if their current visibility settings allow them to be displayed).

Data Processing

The following data processing steps are performed on all incoming data. Each one of them roughly corresponds to a section of the web portal, which visualizes the results of the analyses.

Localization Module

This module allows tracking the geographical origin of the tweet, visualizing which countries are more involved in the discussion about the vaccines.

The geolocation is extracted directly from the tweet, whenever it is possible. Users on Twitter can decide whether to share their location or not at any moment, and whether to geotag the places mentioned in their tweets. If precise geolocation is not available, the module attempts to reconstruct it using the user's "location". This is a free text field located in the user's profile. As such, it may contain imaginative terms or non-existent locations (eg, "over the rainbow" or "the universe"). The module relies on heavy pre-processing, normalization, and cleaning steps to discard most of the noisy locations. The remaining ones are passed to the Google Maps services [26] to determine the most accurate match.

The information is displayed on the web portal as a world map, where countries are shown in different shades of color. The larger the number of tweets coming from that country, the darker the color (the scale is exponential).

Hashtag Analysis

Hashtags are extracted from the most recent tweets only (the last 7 days, updated daily).
We automatically remove a curated selection of hashtags, considered to be of low information content. In particular, we remove all hashtags containing the name of the vaccines that we are

tracking (eg, #pfizer, #moderna, #biontech, etc.), words directly related to the COVID-19 (eg, #covid, #coronavirus, #covidvaccine, etc.) and the ones containing the term "vaccine" only.

Information displayed on our web portal shows the hashtags as a colored treemap, where most tweeted hashtags cover a wider area and are darker in color.

News Sources Analysis

Sensitive topics such as health and vaccinations are fertile ground for the spread of misinformation, as proven by the amount of COVID-19 related fake news which have been debunked in 2020 by fact-checking agencies (eg, PolitiFact [27]) and the precautions taken by the major social networks when dealing with posts mentioning the pandemic (eg, Facebook [28]).

An analysis of the most shared articles is of key importance to understand which sources of information are used by the people to inquire about vaccines.
We run the analysis by collecting all URLs contained in the tweets.
We consider the most recent tweets only (last 7 days, updated daily), to reflect the impact of the most recent news.

URLs are used both in their full form and considering their domain only. Unique URLs and domains are counted and used to provide two different kinds of information: the single most shared webpages (to individuate trending articles) and the most popular sources of information (intended as websites/domains, to individuate the favorite source of information in general).

Factuality Analysis

In order to further investigate the factuality of the URLs shared by the users, we make use of Iffy+ [29], a website that provides an updated list of websites ranked by their factuality level. The lists provided by Iffy are the result of an aggregation of different famous fact-checking websites and trusted sources (eg, FactCheck.org, PolitiFact, and Wikipedia). The list we take into account is composed, for the most part, of websites with a low MBFC Factual level [30] and sources of Fake-News/Misinformation identified by BuzzFeed, FactCheck.org, PolitiFact and Wikipedia. We use this list to perform a factuality analysis over all the collected tweets.

For each URL in a tweet, we check if its domain belongs to one of the websites on the Iffy+ list. If it does, we classify it according to its level of *MBFC factuality* (High, Mixed, Low, Very Low), and its *misinformation category* (eg, Conspiracy, Fake News, ...). Factuality level and misinformation category might be not available for some of the websites ("Not Available"). If a domain is not part of the Iffy+ list, we assume it is a reliable ("Reliable") source of information. All the domains with a factuality level greater than or equal to "High" are labeled as "Reliable". Only 0.0089% of the "Reliable" URLs fall into this category.

We want to highlight that this analysis only explores the reliability of the links that the users are sharing, but not the legitimacy of the tweet as a whole. For example, a user might share a "Fake News" article as a way to joke, mocking it in the text of the tweet. There might also be cases of users sharing links from reliable sources, accompanied by inflammatory or fake captions.

Sentiment Analysis

The sentiment analysis module aims at understanding the attitude of the users when sharing their opinions of the vaccines and their possible side effects. To understand the general sentiment of the crowd when talking about the vaccines, we employ a RoBERTa model [31] trained on tweets and

fine-tuned for the sentiment analysis on the TweetEval Benchmark [32,33]. The model reaches 72.6 +/- 0.4 macro-averaged Recall on the test set.
This kind of module is useful to interpret the general mood of the people that are speaking about the vaccines, about their possible side effects, or even about their vaccination experiences. In particular, this can be very effective to understand if a user is reporting facts, expressing distress, or showing a positive attitude. For each tweet, the sentiment calculated using RoBERTa is then normalized to a discrete set of values (positive, negative or neutral) for ease of visualization.

Our web portal features an interactive line graph to observe how the sentiment varies in time. It allows the visitor to inspect the sentiment globally and compare the trends for the tweets mentioning specific vaccines.

Symptom Extraction

In the last decade people have started discussing about their personal health status on social media more and more often, looking for users with similar experiences, asking for suggestions or reporting unexpected effects after the assumption of medicines. The latter are an interesting type of information, as they might be considered as Adverse Event (AE) indicators for Pharmacovigilance purposes.

Systems for the automatic extraction of ADEs from informal and social media texts are at the core of a growing research trend in the field of Natural Language Processing [34,35]. Moreover, several shared tasks have been recently organized within the ACL community [36,37] to raise interest about this topic.
Our research group studied different combinations of Transformer pretrained models and Conditional Random Fields (CRF) to create an effective deep learning architecture for the task [16]. The best-performing model employs a neural network architecture based on SpanBERT [13] and Conditional Random Fields [38], trained on the Adverse Event Detection dataset of the SMM4H 2019 Shared Task [39]. It represents the current state of the art on the Shared Task [14,15] (Table 2).

Table 2. Performance in terms of Precision (P), Recall (R), and F1 score (F1) of our ADE extraction module against the previous top-performing models on the SMM4H 2019 Shared Task. Data from the public CodaLab leaderboard [40].

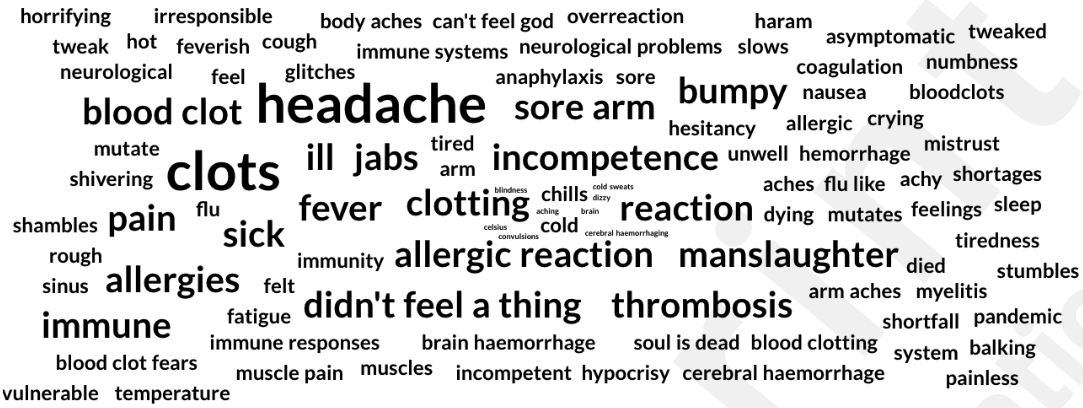
		Relaxed Metrics[a]			Strict Metrics		
		F1	P	R	F1	P	R
Architecture							
	SpanBERT+CRF [15]	**70.2**	60.8	83.0	**46.4**	**39.6**	56.1
	KFU [41]	65.8	55.4	81.0	**46.4**	38.9	**57.9**
	THU_NGN [42]	65.3	**61.4**	69.7	35.6	32.8	38.8
	MIDAS@IIITD [43]	64.1	53.7	79.3	32.8	27.4	40.9
	TMRLeiden [44]	62.5	55.5	71.5	43.1	38.1	49.5

It models the ADE extraction problem as token classification, tagging each word in the text as “inside” or “outside” of a symptom/adverse event.

[a] Relaxed evaluation of the model’s performances. A prediction that does not match exactly the correct AE, but overlaps with it (eg, “headache” instead of “strong headache”) is not discarded but considered as a “partial match” (worth half a point). These evaluation metrics resemble more closely how humans might perceive the correctness of the predictions.

The samples go through five main processing steps: text pre-processing; subword tokenization; BERT modeling; intermediate label prediction; Conditional Random Field; final label aggregation.
The module of our system extracts all the symptoms that are being discussed in the tweets. The data are then aggregated and visualized on the web portal as a Word Cloud. It can be filtered by vaccine and by period of time, to discover what concepts the crowd focused on at different stages of the vaccination campaign.

Figure 2. Possible side effects of the AstraZeneca vaccine, as discussed on Twitter. The Word Cloud is generated using our ADE extraction model and displayed on the web portal. The size of the words is proportional to their frequency.



(Figure 2) shows an example of the Word Cloud calculated using tweets regarding the AstraZeneca vaccine following the thromboembolic events reported in several European countries during March [45].

Model Validation

The Sentiment Analysis and Symptom Extraction modules are based on deep learning models and it is crucial to verify their generalization capabilities outside benchmark environments. To evaluate more rigorously the performance of the modules mentioned above, we sampled and annotated a subset of the collected tweets, to compare the models predictions with human ground-truth labels on real-world data.

A total of 1000 tweets were extracted using stratified sampling to keep the same distribution of tweets over months. Three annotators with high English proficiency (C1) were tasked to mark the sentiment of the tweets on a three-point scale (positive, neutral, negative) and highlight any vaccine-related adverse events mentioned in them.
The gold sentiment of the tweet was decided by majority vote. The gold adverse events of the tweets were decided as: the set of all sequences of words which were highlighted by at least 2 out of 3 annotators. For example, if the annotations were “strong headache”, “headache” and “having a strong headache”, the final annotation would be “headache”.

The human-generated annotations are used as ground-truth to evaluate the performance of the two deep learning modules on the real-world data and compare them with their performance on the benchmark datasets. Results are discussed in the section “Results - Overall Results - Model Validation”.

Results

The following sections present: (i) some observations on the overall results produced by the different modules of the dashboard; (ii) an example of an in-depth analysis focused on the AstraZeneca vaccine based on the collected data. The analyses are performed on data relating to the period December 10th, 2020 - September 7th, 2021.

Overall Results

First of all, we performed an initial analysis on the number of unique tweets and unique user accounts present in the collected data. As mentioned in Methods - Data Collection, we took some precautions to avoid collecting duplicated data or skewing the dataset by giving more weight to tweets posted by popular accounts. To verify if these strategies were successful, we inspect the ratio of unique tweets and users in the dataset, month-by-month and overall.

(Figure 3) shows the distribution of users depending on how many times their tweets appear in the dataset. We can clearly see a long-tail distribution, where 75% of the users only tweeted once, 92% of users tweeted at most three times and 98% of users tweeted at most 10 times (ie, on average once per month). Looking at the users that tweeted more, most of them were news outlets, who tweeted from 50 to 578 times in the considered timespan (0.18% of the total users). The long-tail distribution is a good sign, as it shows that most of the users from which we collect tweets are likely regular users and not influencers or content-farms.

We then look at the origin of the tweets that compose the dataset. (Figure 4) shows that 95% of the total tweets were posted by users that tweeted less than 100 times in the considered timeframe. This is another positive indication that the collection of tweets is not heavily influenced by a small number of super-accounts, and thus the subsequent analysis should not suffer from this kind of bias.

Figure 3. Distribution of users depending on how many times they tweeted (y axes in logarithmic scale).

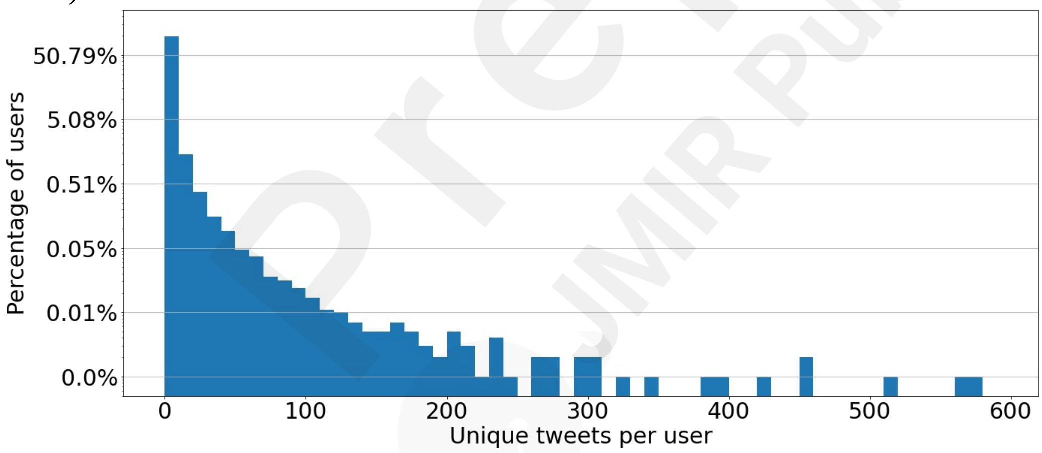
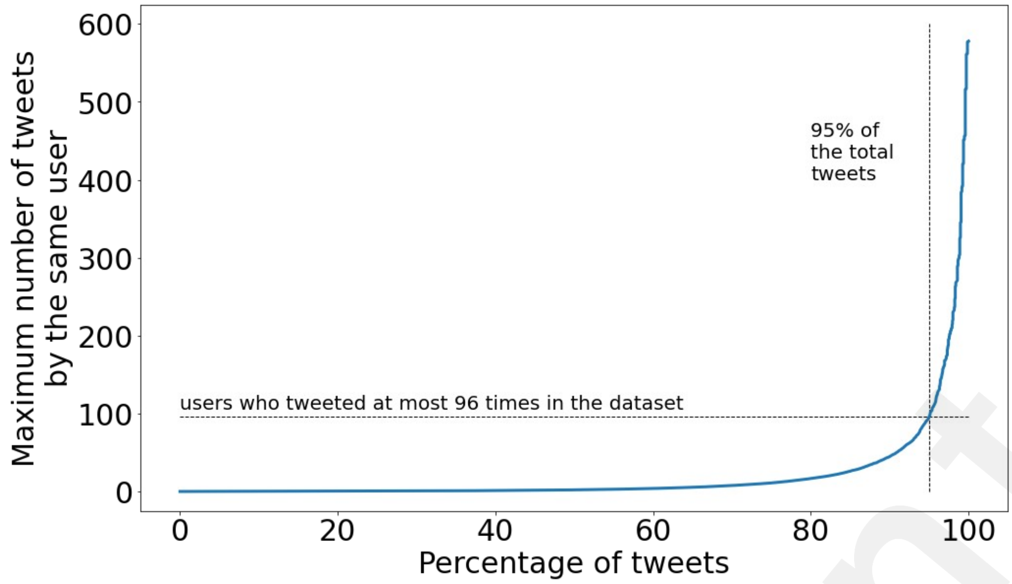


Figure 4. Percentage of tweets produced by a group of users, depending on how many tweets the user produced. 95% of the tweets in the dataset are produced by users who tweeted at most 96 times in the considered timespan.



Finally, we calculated some statistics on a monthly basis and reported them in (Table 3). The mode and median are 1, confirming the findings discussed above. The average number of tweets per user remained stable at around 1.4 during the first months (December 2020 - March 2021). It then increased to 1.5 in the period between April and June, following the start of the vaccination campaigns and the AstraZeneca controversy (probably due to heightened news coverage). Following June, the average number of tweets per user went down again.
The number of unique tweets and unique users considered each month is roughly stable.

Table 3. Statistics on the unique number of tweets and users for each month in the collected data.

				tweets per user				
		unique tweets (n)	unique users (n)	max	avg	std	mode	median
Month								
	2020-12[b]	21235	15983	40	1.32	1.29	1	1
	2021-01	42891	30294	71	1.42	1.76	1	1
	2021-02	36897	25102	98	1.47	1.98	1	1
	2021-03	51469	35402	181	1.45	2.47	1	1
	2021-04	62697	41160	117	1.52	2.45	1	1
	2021-05	48785	32263	134	1.51	2.45	1	1
	2021-06	41364	27397	154	1.51	2.45	1	1
	2021-07	42742	29371	139	1.46	2.26	1	1
	2021-08	41596	29942	232	1.39	2.09	1	1
	2021-09[b]	7064	5833	27	1.21	0.84	1	1
	all	396740	196011	578	2.02	6.19	1	1

Localization

Since we are only considering English-language tweets, the most active countries are the USA, Canada, and the United Kingdom, followed by Nigeria, India, Australia and finally various European countries. Despite the language limitation that we imposed, the system detected tweets from almost

[b] Partial data, does not span the entirety of the month.

all the countries in the world.
We plan to remove the language limitation in the near future, by means of the usage of automated translation services.

Hashtags

Most of the top hashtags are related to the concepts of “health”, “news” or mention specific countries that made it to the top headlines due to recent outbreaks and similar accidents.

News Sources

The current data show a reassuring trend: the most popular sources of information are renowned newspapers (such as The New York Times or The Guardian), official institutional websites (eg, www.gov.uk), and scientific authorities (eg, the EMA and WHO). It is also interesting to note that, since the monitoring started in December 2020, the video-sharing platform YouTube has always been among the top-15 most shared domains. The top-5 most shared articles are displayed on the website as clickable links (displaying the URL and title of the page), while the 15 most popular domains are shown as a bar graph.

Factuality

The vast majority of the shared URLs are classified as having a “Reliable” level of factuality (98%, see (Figure 5)). This seems to be confirmed if we look at the five most shared domains: theguardian.com (3.22%), nytimes.com (2.75%), reuters.com (2.40%), cnbc.com (1.77%), abc.net.au (1.56%).

The remaining 2% is composed by domains classified mostly as *Low* and *Mixed* (ie, a website that is known to share both factual and non-factual information). (Figure 6) shows the factuality distribution of “Unreliable” URLs (pay attention to the logarithmic scale).

Looking at the misinformation categories for “Unreliable” domains (Figure 7), we can see that 49% are classified as “Conspiracy-Pseudoscience”, 49% as generic “Fake-News” sources, and the remaining are subject to political biases.

Figure 5. All the plots are in logarithmic scale. Percentage of the Reliable and Unreliable URLs shared.

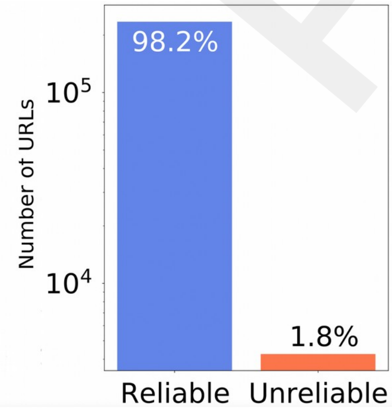


Figure 6. Distribution of *MBFC misinformation categories* for “Unreliable” URLs: Conspiracy-Pseudoscience (CP), Fake-News (FN), Not Available (N/A), Right-Center bias (RC), Right bias (R) and Left bias (L). y-axis in logarithmic scale.

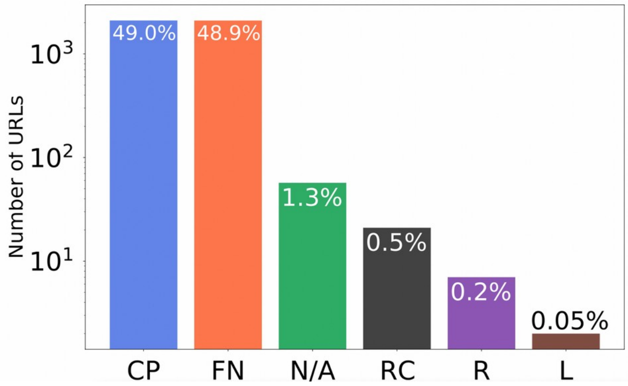
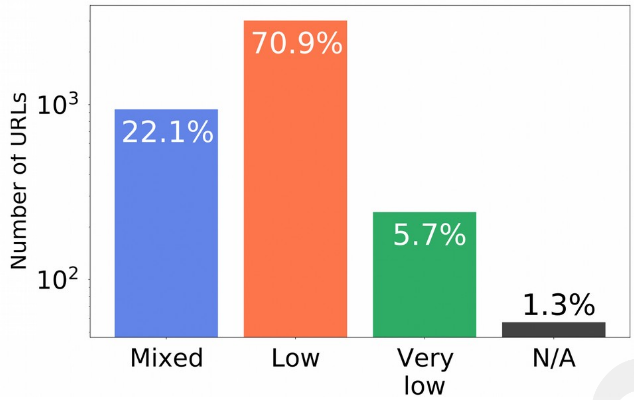


Figure 7. Distribution of *MBFC factuality level* for “Unreliable” URLs. y-axis in logarithmic scale.



Sentiment Analysis

We observed that the global sentiment of the tweets we analyzed is neutral/negative, for most of the period of observation, with occasional spikes of positivity for individual vaccines. The negative trend might be enhanced by the fact that shocking, controversial or tragic news tend to be shared and spread more easily on the internet, when compared to other kinds of news.

Symptom Extraction

In the days preceding March 11, the most prominent concepts in AstraZeneca's Word Cloud were “headache” and “fever” but, as soon as thromboembolic events started being discussed on the internet, the system detected the shift in topic and words such as “clots” and “thrombosis” quickly became noticeable in the Cloud.
As regards the other two vaccines, “allergic reactions”, “headache” and “fever” have always been among the most shared and discussed adverse events. “Anaphylaxis” was one of the major concepts on Pfizer-BioNTech's Cloud for a long period of time at the beginning of the vaccination campaign but is now slowly losing traction (you can see this in the Word Cloud on our web portal).

This model is able to identify tweets containing potential adverse events and to highlight the mention of the symptoms. However, there are no mechanisms in place to verify the reliability of the tweets and there is no human fact-checking involved in the process. This means that, for the time being, there is virtually no distinction between symptoms which were actually reported by the users and exaggerations or hoaxes. This limitation is clearly stated on the web portal and the viewers are encouraged to further inspect the tweets on their own to have a clearer idea of what kind of messages lead to the prediction of the extracted symptoms. Clicking on any word in the Word Cloud displays a selection of the analyzed tweets that mentioned that concept in the selected time period.

Section "Evolution of mentioned symptoms over time" contains an analysis of the information that can be extracted by the representations produced by this module.

Finally, we would like to recall that the system was trained solely on the data provided during the SMM4H 2019 Shared Task. Even though it is one of the best performing models on this task, it still suffers from the limitations of current ADE extraction systems, such as the difficulty in making reliable distinctions between side effects (caused by medications), symptoms (caused by illnesses), and the names or descriptions of some medical conditions. For example, in the sentence "I have a *slipped vertebrae* and a *degenerative disk*.", the two medical conditions are identified as side effects by the system.
This is a common problem for such systems, which are often trained on datasets that are limited in size and linguistic variety.

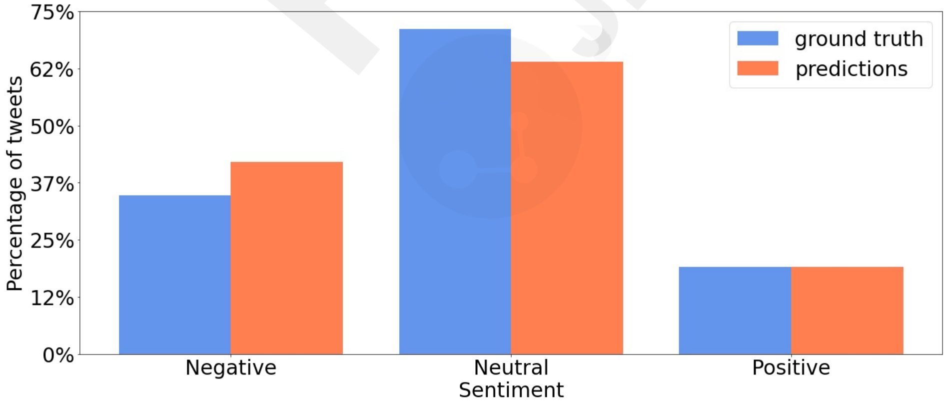
Model Validation

We experimentally evaluated the performance of both the Sentiment and Symptom Extraction modules using the subset of 1000 manually-annotated tweets we created.

The performance of the sentiment module on the real data is in line with the one on the benchmark dataset, and its predictions are close to the ground truth. (Figure 8) shows the sentiment distribution of the ground truth labels (blue) and the predictions of the model (orange). The model leans slightly more towards Negative sentiment. The performance (macro-averaged Recall) on the subset of our data was 72.1. It shows excellent generalization capabilities, and it is in line with the performance recorded on the benchmark dataset of 72.6 +/- 0.4.

To evaluate the Symptom Extraction module, we sampled our dataset to have the same ratio of AE/noAE tweets as the benchmark dataset SMM4H (57:43). The obtained relaxed F1 score is 63.3 +/- 0.7 (average over 10 sampling procedures), against 70.2 recorded on SMM4H. The gap in performance may be caused by the difference in the types of adverse events present in the two datasets. For example, the benchmark dataset focuses on sleep disorders and weight gain/loss, while the data we collected contain more instances of arm soreness and blood clotting, which the model has never encountered during training.

Figure 8. Comparison of the sentiment distributions of the manually-annotated ground truth labels (blue) and the model predictions (orange).



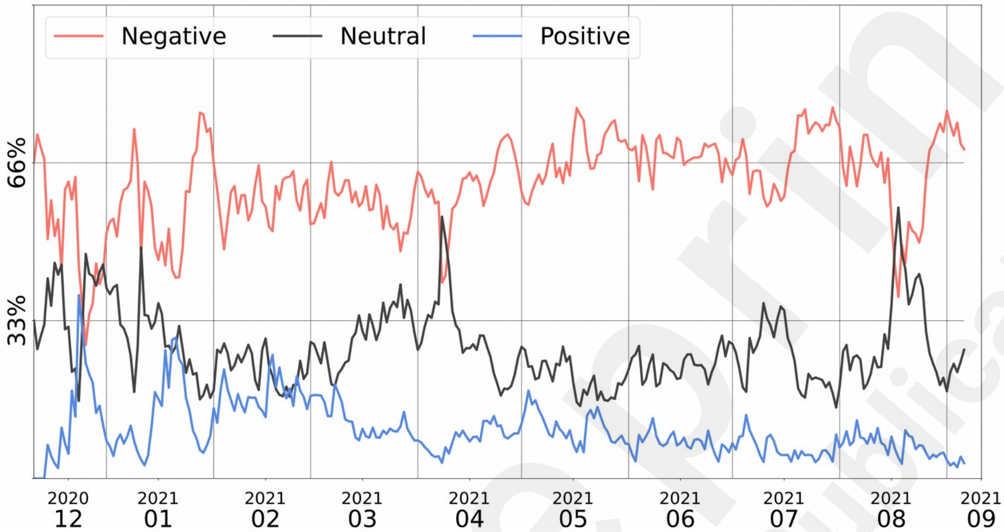
Case Study: AstraZeneca

To demonstrate the possible uses of our monitoring system as a research tool, we created a brief report regarding the AstraZeneca vaccine. In particular, we focused on analyzing the phenomenon of the alleged correlation between the vaccine and some specific side effects (eg, blood clots), comparing it with the other monitored vaccines.

Sentiment Trends for AstraZeneca

We start by having a general overview of the sentiment of the crowd toward the vaccine, and how it varied in time. (Figure 9) shows the day-by-day percentage of positive, neutral, and negative tweets about AstraZeneca, from the day the monitoring started (11 December 2020) to the most recent date at the time of writing (early September 2021).

Figure 9. Monthly sentiment distribution in AstraZeneca-related tweets. The y-axis represents the percentage of negative (top, orange), neutral (middle, grey), and positive (bottom, blue) sentiment in the analyzed tweets. It is clear that the prevalent sentiment overall is “negative”, but we can observe spikes of non-negativity in December and January.



We can see the sentiment towards the vaccine is mostly negative for the whole time period. This is probably due to the tendency of negative and worrying topics or critical opinions to spread more easily on the Internet. Approximately one third of the tweets is neutral, corresponding to people sharing factual information about the vaccine or showing neutrality and detachment towards the topic.

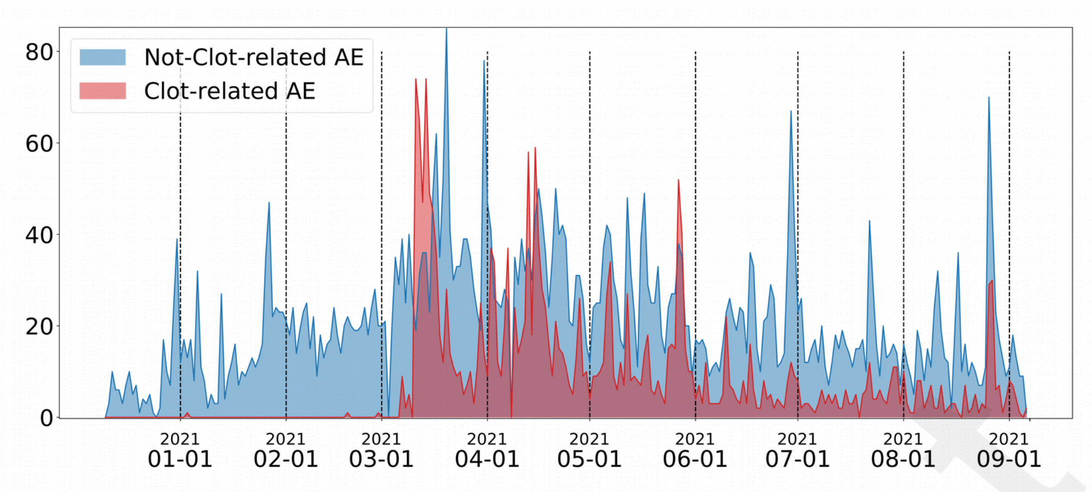
There is a noticeable trend of “non-negativity” between December and January, when positive and neutral tweets cover more than half of the discussion.

This might be related to the publication of an important study [46] about the efficacy of AstraZeneca and its approval by the EMA.

Mentions of Thromboembolic Events

We then compared the frequency with which Twitter users mentioned adverse events related to “thrombosis” and “blood clotting” compared to other vaccine side effects.

Figure 10. Number of tweets mentioning clot-related and not-clot-related keywords for AstraZeneca (time on the x-axis, number of tweets on the y-axis). The number of tweets mentioning clot-related AEs was initially next to zero, spiked in March 2021 due to media coverage but has been gradually diminishing ever since. Tweets mentioning non-clot-related AEs show a more stable trend over time.



(Figure 10) shows for each day the number of detected tweets that contained clot-related AEs (red series) and any other adverse event (blue series).

We can see that the absolute number of tweets discussing AstraZeneca and its adverse events increased from December 2020 to February 2021, but blood clotting events were rarely discussed on Twitter.

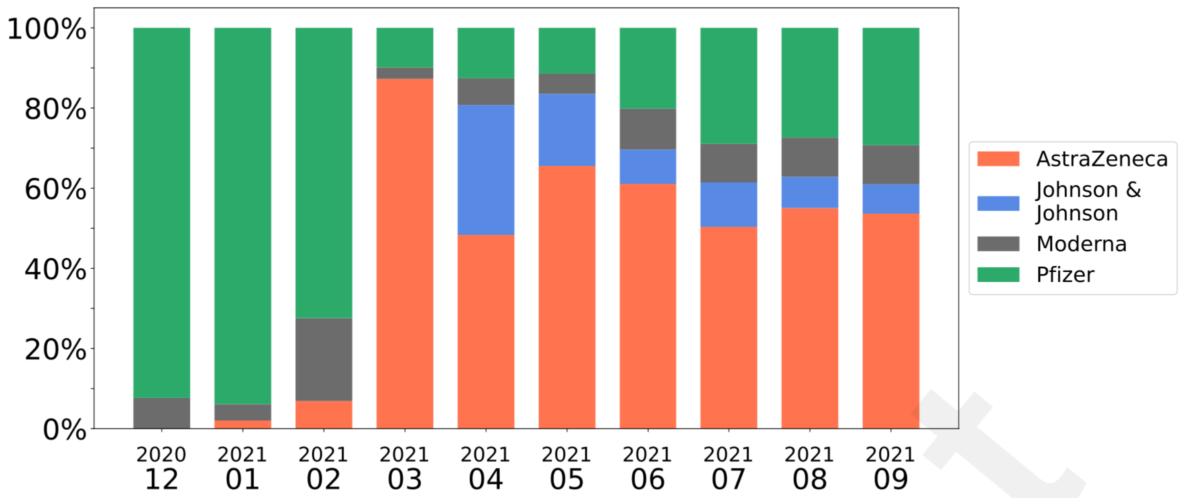
This changed in the first half of March 2021, when the number of tweets discussing clot-related AEs had a peak. At that time, some European states (eg, Germany) stopped the inoculations of the AstraZeneca vaccine due to the possible correlation between the clots and the vaccine and some suspicious deaths from ischemia.

Since then, the public attention on clot-related AEs remained high and peaked periodically (see the red series), without losing track of the other topics (the number of tweets discussing other AEs remains high).

As specified above, not all the tweets with clots-related references are AE reports: most of them come from people sharing or commenting news pieces about the vaccine.

We can also observe that in the last month the chatter about AstraZeneca has diminished, as the blue and red series report less than 20 tweets per day.

Figure 11. Monthly distribution of vaccine names mentioned in tweets with clot-related keywords (time on the x-axis, percentage of tweets on the y-axis). Most of the tweets were discussing clot-related AEs connected to the Pfizer vaccine before March 2021, when the focus suddenly shifted to AstraZeneca.



(Figure 11) offers a different perspective on the phenomenon: we collected all tweets mentioning blood clots and thrombosis and divided them according to which vaccines they deal with. Before March 2021, most of the tweets dealing with clot-related AE were relative to the Pfizer vaccine (75-85%). With the wide news coverage about the cases related to AstraZeneca, the trend changed drastically, and over 80% of the tweets mentioning this kind of events was discussing AstraZeneca.

Evolution of Mentioned Symptoms over Time

The wide news coverage had a strong influence on the topics of discussion among the Twitter users. This can be seen even more clearly in (Figure 12). It shows three series of Word Clouds that represent how the main topics discussed on Twitter varied in time. The first row shows the most frequent AEs globally discussed (considering all tweets), month-by-month. The following rows show the evolution of the topics for the tweets that mention AstraZeneca, Moderna or Pfizer only.

Figure 12. Evolution of the global Word Cloud (top row, all vaccines included) and the specific Word Cloud of the following vaccines: AstraZeneca, Johnson&Johnson, Moderna, Pfizer. The suspected ADEs are extracted using our model.

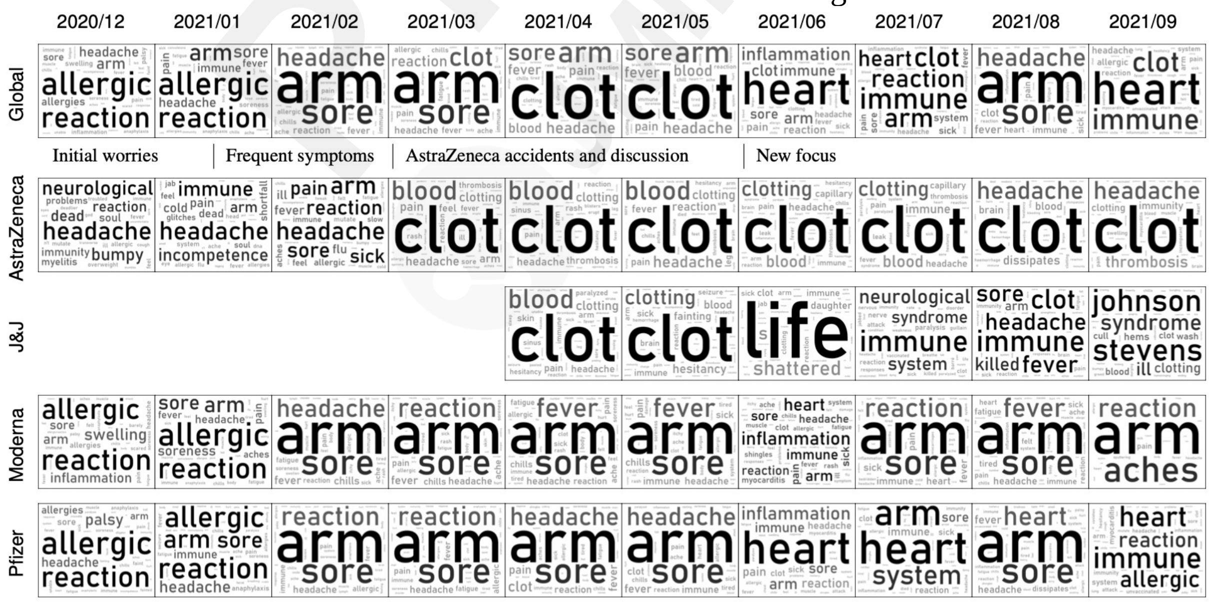
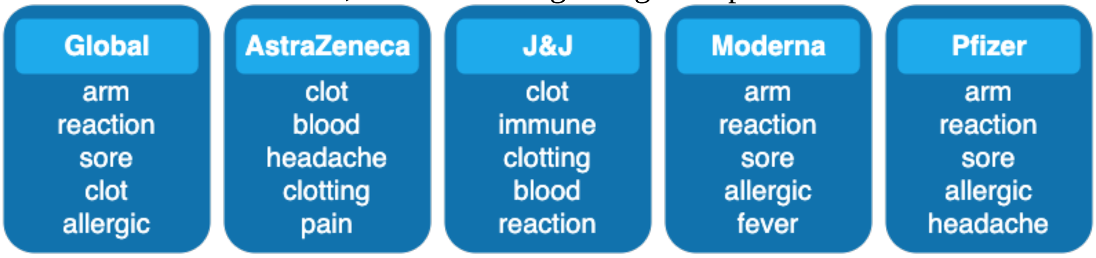


Figure 13. Top-5 most frequently mentioned terms globally and for the following vaccines:

AstraZeneca, Johnson&Johnson, Moderna, Pfizer. This takes into consideration all the collected tweets from December 15, 2020 to the beginning of September 2021.



We can see that in the first two months (2020-12 to 2020-01) all the discussions were focused on widespread worries and doubts of the users (eg, allergies, neurological problems, immune responses etc.).

During the following months, as the vaccination campaign proceeded, the focus slowly shifted towards the most common side effects that the vaccinated population was experiencing (eg, soreness at the arm, feeling sick, headache).

The news about AstraZeneca on March caused a dramatic shift of topic not only in the tweets regarding that particular vaccine, but also globally: the word "clot" suddenly appears in the global Word Cloud and becomes the most discussed topic for the following months (it also influences Pfizer's Word Cloud, where the "clot" topic becomes slightly visible in April).

Looking at the latest available data, we can see that "blood clots" are still the most trending topic for AstraZeneca, but the global discussion has finally moved towards other topics, such as "heart" problems. That said, if we look at all of the collected data, from December 2020 to September 2021, we can see how "clot" is the fourth most mentioned term globally (Figure 13), surpassed in popularity only by the broader concepts "arm", "reaction" and "sore". This shows how great of an impact this episode had on the social media.

Discussion

Intended Use Cases

Our web portal could be useful to different categories of users:

- *General public.*
  Thanks to the intuitive interface and graphics, generic users can keep themselves up-to-date and be made aware of the kind of news which are circulating, what symptoms are being discussed for the various vaccines and under which terms.
- *Journalists and news outlets.*
  The section of the web portal dedicated to news trends might provide insights for the press to better understand the digital audience and help in fighting misinformation. The other information might be interesting to explore to discover the latest most discussed topics.
- *People in the healthcare sector.*
  The information on the most shared symptoms and possible adverse events might be helpful to point the attention of the experts towards particular effects of the new vaccines.
- *Scholars working in biomedical Natural Language Processing.*
  The code of the ADE extraction architecture is publicly available, and the web portal includes an explanatory page about the various implemented modules. The objective is to raise the interest of the NLP community on this topic and open the door to suggestions and possible collaborations.

Limitations

This project collects data from user-generated, unfiltered content and makes use of automatic tools that have low and no human supervision. Therefore, it is important to highlight some limiting factors:

- Language barrier. As stated in the first sections, the current system is only able to analyze texts written in English. The COVID-19 vaccines are being distributed and discussed in several non-english-speaking countries, so this dataset is only a partial representation of public opinion. As stated in the Data Collection section, we plan to overcome this limitation with the use of multilingual models and/or automated translation services. We are already collecting tweets in other languages for the same time-period and this will allow us to perform a complete comparative analysis in the future.
- Demographics of Twitter users. Twitter is often used as a means to understand and monitor crowd opinions and real-world phenomena. However, it is not always the case that Twitter users are a representative sample of the population we are actually interested in. A population can be examined along various axes (e.g., age, geography, gender, ethnicity) and specific social media environments tend to overrepresent some sets of the population (e.g., users coming from densely populated areas, higher level of education, higher income or computer literacy) [47,48].
- Bias, echo chambers and misinformation spread on social media. Social media are also infamous for the creation of echo chambers [49], where users of the same mindset end up aggregating. This can “artificially” increase engagement with polarizing posts, which in turn become more visible and gain more weight in the analyses. Social media are highly polarizing environments, in which shocking, controversial and generally “negative” posts are rewarded (and therefore can be found more frequently in the collected data) [50,51].
  Our system tries to cope with this by handling data deduplication (removing viral copy-pasted tweets) and collecting the most recent tweets (as opposed to the most popular). This, however, does not remove the threats of echo chambers and misinformation. As a future work, we plan to add a new module based on our previous work [52] to better analyze phenomena related to the spread of misinformation.
- Correctness of deep learning modules. Both the sentiment analysis and the symptom extraction module are machine learning modules, and as such can perform prediction errors with a known probability. If the data are shown to the public, users must be aware that they have to be taken with a grain of salt. This is why, on our dashboard, we make sure to include a disclaimer to warn the user about this issue whenever we display data produced by ML algorithms.

Conclusions

We presented a tool connected with a web portal to monitor and display some key aspects of the public's reaction to COVID-2019 vaccines.

The idea was born from the awareness that, in the current phase of the pandemic, it is of key importance to create tools to monitor reactions, opinions, doubts and feedback of the population on the vaccines. Social media are a precious source of raw information, which can be exploited to gain insights for pharmacovigilance purposes (guiding the attention of healthcare experts on emerging effects) and help in fighting misinformation.

The system also provides an overview of the opinions of the Twittersphere through graphic representations, to make them accessible to different categories of users.

One of the main features of this tool is the extraction of suspected adverse events from tweets with a Deep Learning model, which proved to be reactive to the shifts of topic in the Internet chatter.
A future improvement could be the extraction of ADE from tweets of different languages, using a multi-lingual model or an automated translation service.

All code, tweet ids and the pre-computed statistics of the collected tweets are available at https://github.com/AilabUdineGit/covid-vaccines-tools.

Conflicts of Interest

None declared.

Abbreviations

AE: adverse event
ADE: adverse drug event
ADR: adverse drug reaction
ML: machine learning
URL: uniform resource locator

References

1. Forni G, Mantovani A, Forni G, Mantovani A, Moretta L, Rappuoli R, Rezza G, Bagnasco A, Barsacchi G, Bussolati G, Cacciari M, Cappuccinelli P, Cheli E, Guarini R, Bacci ML, Mancini M, Marcuzzo C, Morrone MC, Parisi G, Pasquino G, Patrono C, Curzio AQ, Remuzzi G, Roncaglia A, Schiaffino S, Vineis P. COVID-19 vaccines: where we stand and challenges ahead. Cell Death Differ Nature Publishing Group; 2021;28(2):626–639. PMID:33479399
2. Goel RK, Nelson MA, Goel VY. COVID-19 vaccine rollout—scale and speed carry different implications for corruption. J Policy Model 2021;43(3):503–520. PMID:33967361
3. Sah P, Vilches TN, Moghadas SM, Fitzpatrick MC, Singer BH, Hotez PJ, Galvani AP. Accelerated vaccine rollout is imperative to mitigate highly transmissible COVID-19 variants. EClinicalMedicine [Internet] 2021;35:100865. doi: 10.1016/j.eclinm.2021.100865
4. Kruikemeier S. How political candidates use Twitter and the impact on votes. Comput Human Behav 2014;34:131–139. doi: 10.1016/j.chb.2014.01.025
5. Jungherr A. Twitter use in election campaigns: A systematic literature review. J Inf Technol Polit Routledge; 2016;13(1):72–91. doi: 10.1080/19331681.2015.1132401
6. Cinelli M, Quattrociocchi W, Galeazzi A, Valensise CM, Brugnoli E, Schmidt AL, Zola P, Zollo F, Scala A. The COVID-19 social media infodemic. Sci Rep Nature Publishing Group; 2020;10(1):1–10. PMID:33024152
7. Sharevski F, Huff A, Jachim P, Pieroni E. (Mis)perceptions and engagement on Twitter: COVID-19 vaccine rumors on efficacy and mass immunization effort. Int J Inf Manag Data Insights; 2022;2(1):100059. doi: https://doi.org/10.1016/j.jjimei.2022.100059
8. Karafillakis E, Martin S, Simas C, Olsson K, Takacs J, Dada S, Larson HJ. Methods for social media monitoring related to vaccination: Systematic scoping review. JMIR Public Heal Surveill 2021 Feb;7(2):e17149. PMID:33555267
9. Yan C, Law M, Nguyen S, Cheung J, Kong J. Comparing public sentiment toward COVID-19 vaccines across Canadian cities: Analysis of comments on reddit. J Med Internet Res 2021

Sep;23(9):e32685. PMID:34519654
10. Kwok SWH, Vadde SK, Wang G. Tweet topics and sentiments relating to COVID-19 vaccination among Australian twitter users: Machine learning analysis. J Med Internet Res 2021 May;23(5):e26953. PMID:33886492
11. Kummervold PE, Martin S, Dada S, Kilich E, Denny C, Paterson P, Larson HJ. Categorizing vaccine confidence with a transformer-based machine learning model: Analysis of nuances of vaccine sentiment in twitter discourse. JMIR Med Informatics 2021 Oct;9(10):e29584. doi: 10.2196/29584
12. Santus E, Marino N, Cirillo D, Chersoni E, Montagud A, Chadha AS, Valencia A, Hughes K, Lindvall C. Artificial intelligence-aided precision medicine for COVID-19: Strategic areas of research and development. J Med Internet Res 2021;23(3). PMID:33560998
13. Joshi M, Chen D, Liu Y, Weld DS, Zettlemoyer L, Levy O. Spanbert: Improving pre-training by representing and predicting spans. Trans Assoc Comput Linguist 2020;8:64–77. doi: 10.1162/tacl_a_00300
14. Portelli B, Passabì D, Lenzi E, Serra G, Santus E, Chersoni E. Improving Adverse Drug Event Extraction with SpanBERT on Different Text Typologies. arXiv Prepr arXiv210508882 2021; Available from: http://arxiv.org/abs/2105.08882
15. Portelli B, Lenzi E, Chersoni E, Serra G, Santus E. BERT Prescriptions to avoid unwanted headaches: A comparison of transformer architectures for adverse drug event detection. EACL 2021 - 16th Conf Eur Chapter Assoc Comput Linguist Proc Conf 2021. p. 1740–1747. doi: 10.18653/v1/2021.eacl-main.149
16. Scaboro S, Portelli B, Chersoni E, Santus E, Serra G. NADE: A Benchmark for Robust Adverse Drug Events Extraction in Face of Negations. Proc Seventh Work Noisy User-generated Text (W-NUT 2021) Online: Association for Computational Linguistics; 2021. p. 230–237. doi: 10.18653/v1/2021.wnut-1.26
17. Lynch CJ, Gore R. Short-Range Forecasting of COVID-19 During Early Onset at County, Health District, and State Geographic Levels Using Seven Methods: Comparative Forecasting Study. J Med Internet Res; 2021 Mar;23(3):e24925. doi: 10.2196/24925
18. Benis A, Chatsubi A, Levner E, Ashkenazi S. Change in Threads on Twitter Regarding Influenza, Vaccines, and Vaccination During the COVID-19 Pandemic: Artificial Intelligence--Based Infodemiology Study. JMIR Infodemiology; 2021 Oct;1(1):e31983. doi: 10.2196/31983
19. Zhang J, Wang Y, Shi M, Wang X. Factors Driving the Popularity and Virality of COVID-19 Vaccine Discourse on Twitter: Text Mining and Data Visualization Study. JMIR Public Heal Surveill; 2021 Dec;7(12):e32814. doi: 10.2196/32814
20. Liew TM, Lee CS. Examining the Utility of Social Media in COVID-19 Vaccination: Unsupervised Learning of 672,133 Twitter Posts. JMIR public Heal Surveill JMIR Publications; 2021 Nov 3;7(11):e29789–e29789. doi: 10.2196/29789
21. Muric G, Wu Y, Ferrara E. COVID-19 Vaccine Hesitancy on Social Media: Building a Public Twitter Data Set of Antivaccine Content, Vaccine Misinformation, and Conspiracies. JMIR Public Heal Surveill 2021 Nov;7(11):e30642. doi: 10.2196/30642
22. Martínez Beltrán E, Quiles M, Pastor-Galindo J, Nespoli P, García Clemente FJ, Gomez Marmol F. COnVIDa: COVID-19 multidisciplinary data collection and dashboard. J Biomed Inform 2021;117:103760. doi: 10.1016/j.jbi.2021.103760
23. DeVerna MR, Pierri F, Truong BT, Bollenbacher J, Axelrod D, Loynes N, Torres-Lugo C, Yang K-C, Menczer F, Bryden J. CoVaxxy: A Collection of English-language Twitter Posts About COVID-19 Vaccines. 2021; Available from: http://arxiv.org/abs/2101.07694
24. Sharma K, Seo S, Meng C, Rambhatla S, Liu Y. COVID-19 on Social Media: Analyzing Misinformation in Twitter Conversations. 2020.
25. Twitter Inc., Twitter API Documentation, 2021. URL: https://developer.twitter.com/en/docs/twitter-api [accessed: Nov 22, 2021]

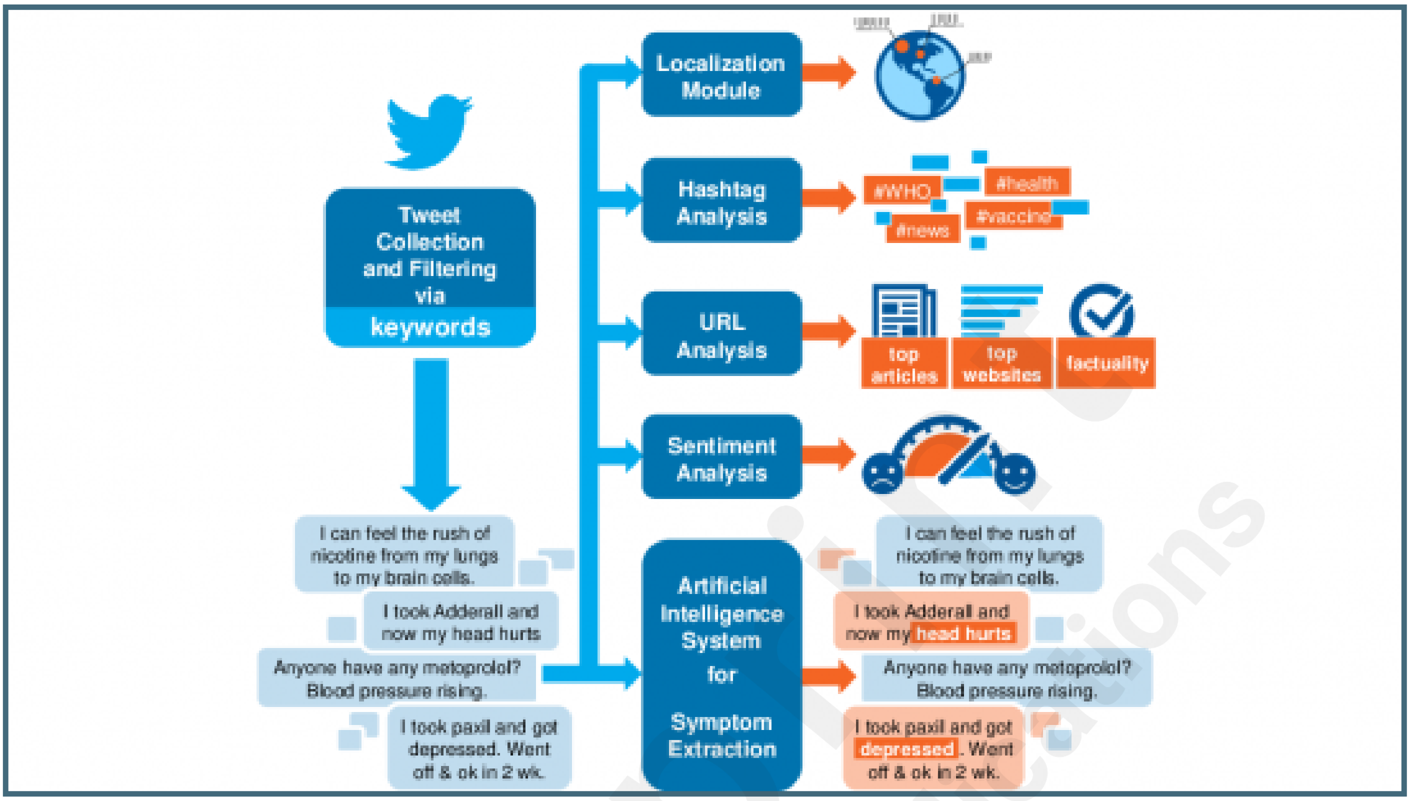
26. Google, Google Maps Platform, 2021. URL: https://developers.google.com/maps [accessed: Nov 22, 2021]
27. Poynter Institute, PolitiFact Coronavirus, 2020. URL: https://www.politifact.com/coronavirus/ [accessed: Apr 18, 2021]
28. Facebook, Keeping People Safe and Informed About the Coronavirus, 2020. URL: https://about.fb.com/news/2020/12/coronavirus/ [accessed: Gen 13, 2021]
29. Iffy.news, Iffy+ Mis/Disinfo Sites, 2021. URL: https://iffy.news/iffy-plus/ [accessed: Sep 20, 2021]
30. Media Bias Fact Check LLC, Media Bias/Fact Check, 2021. URL: https://mediabiasfactcheck.com [accessed: Sep 20, 2021]
31. Liu Y, Ott M, Goyal N, Du J, Joshi M, Chen D, Levy O, Lewis M, Zettlemoyer L, Stoyanov V. RoBERTa: A Robustly Optimized BERT Pretraining Approach. arXiv Prepr arXiv190711692 2019; Available from: http://arxiv.org/abs/1907.11692
32. Rosenthal S, Farra N, Nakov P. SemEval-2017 Task 4: Sentiment Analysis in Twitter. Proc SemEval 2017.
33. Barbieri F, Camacho-Collados J, Espinosa Anke L, Neves L. TweetEval: Unified Benchmark and Comparative Evaluation for Tweet Classification. Proc Find EMNLP 2020. p. 1644–1650. doi: 10.18653/v1/2020.findings-emnlp.148
34. Karimi S, Wang C, Metke-Jimenez A, Gaire R, Paris C. Text and data mining techniques in adverse drug reaction detection. ACM Comput Surv ACM New York, NY, USA; 2015;47(4):1–39. doi: 10.1145/2719920
35. Sarker A, Gonzalez G. Portable automatic text classification for adverse drug reaction detection via multi-corpus training. J Biomed Inform 2015;53:196–207. PMID:25451103
36. Weissenbacher D, Sarker A, Paul MJ, Gonzalez-Hernandez G. Overview of the Third Social Media Mining for Health (SMM4H) Shared Tasks at EMNLP 2018. Proc 2018 EMNLP Work SMM4H 3rd Soc Media Min Heal Appl Work & Shar Task Brussels, Belgium: Association for Computational Linguistics; 2019. p. 13–16. doi: 10.18653/v1/w18-5904
37. Weissenbacher D, Sarker A, Magge A, Daughton A, O'Connor K, Paul MJ, Gonzalez-Hernandez G. Overview of the Fourth Social Media Mining for Health (SMM4H) Shared Tasks at ACL 2019. Proc ACL Soc Media Min Heal Appl Work & Shar Task 2019. p. 21–30. doi: 10.18653/v1/w19-3203
38. Lafferty J, Mccallum A, Pereira F. Conditional Random Fields : Probabilistic Models for Segmenting and Labeling Sequence Data Abstract. Proc ICML 1999. p. 282–289.
39. HLP @ UPenn IBI, Social Media Mining for Health Applications (SMM4H) Workshop & Shared Task 2019, 2019. URL: https://healthlanguageprocessing.org/smm4h19/challenge [accessed: Mar 22, 2020]
40. CodaLab, CodaLab Competition – SMM4H'19 – Shared Task, 2020. URL: https://competitions.codalab.org/competitions/20798#results (Sub-Task 2, ADR extraction, Post-Evaluation) [accessed: Nov 02, 2020]
41. Miftahutdinov Z, Alimova I, Tutubalina E. KFU NLP Team at SMM4H 2019 Tasks: Want to Extract Adverse Drugs Reactions from Tweets? BERT to The Rescue. Proc ACL Work Soc Media Min Heal Appl 2019. p. 52–57. doi: 10.18653/v1/w19-3207
42. Ge S, Qi T, Wu C, Huang Y. Detecting and Extracting of Adverse Drug Reaction Mentioning Tweets with Multi-Head Self Attention. Proc ACL Soc Media Min Heal Appl (# SMM4H) Work & Shar Task 2019. p. 96–98. doi: 10.18653/v1/w19-3214
43. Mahata D, Anand S, Zhang H, Shahid S, Mehnaz L, Kumar Y, Shah RR. MIDAS@SMM4H-2019: Identifying Adverse Drug Reactions and Personal Health Experience Mentions from Twitter. Proc ACL Work Soc Media Min Heal Appl 2019. p. 127–132. doi: 10.18653/v1/w19-3223
44. Dirkson A, Verberne S. Transfer Learning for Health-related Twitter Data. Proc ACL Work Soc

Media Min Heal Appl 2019. p. 89–92. doi: 10.18653/v1/w19-3212
45. European Medicines Agency, COVID-19 Vaccine AstraZeneca: PRAC investigating cases of thromboembolic events - vaccine's benefits currently still outweigh risks - Update, 11/03/2021. URL: https://www.ema.europa.eu/en/news/covid-19-vaccine-astrazeneca-prac-investigating-cases-thromboembolic-events-vaccines-benefits/ [accessed: Jun 14, 2021]
46. Knoll MD, Wonodi C. Oxford–AstraZeneca COVID-19 vaccine efficacy. Lancet 2021;397(10269):72–74. PMID:33306990
47. Smith A, Brenner J. Twitter Use 2012. PewResearchCenter. May 31,2012. https://www.pewresearch.org/internet/2012/05/31/twitter-use-2012/ [accessed: Gen 14, 2022].
48. Mislove A, Lehmann S, Ahn Y-Y, Onnela J-P, Rosenquist J. Understanding the Demographics of Twitter Users. Proc Int AAAI Conf Web Soc Media 2021;5(1):554–557. Available from: https://ojs.aaai.org/index.php/ICWSM/article/view/14168
49. Usher N, Holcomb J, Littman J. Twitter Makes It Worse: Political Journalists, Gendered Echo Chambers, and the Amplification of Gender Bias. Int J Press 2018;23(3):324–344. doi: 10.1177/1940161218781254
50. Fine J, Hunt M. Negativity and Elite Message Diffusion on Social Media. Polit Behav 2021;1–19. doi: 10.1007/s11109-021-09740-8
51. Buder J, Rabl L, Feiks M, Badermann M, Zurstiege G. Does Negatively Toned Language Use on Social Media Lead to Attitude Polarization? 2020. doi: 10.31234/osf.io/dx9ws
52. Portelli B, Zhao J, Schuster T, Serra G, Santus E. Distilling the Evidence to Augment Fact Verification Models. Proc Third Work Fact Extr Verif Online: Association for Computational Linguistics; 2020. p. 47–51. doi: 10.18653/v1/2020.fever-1.7

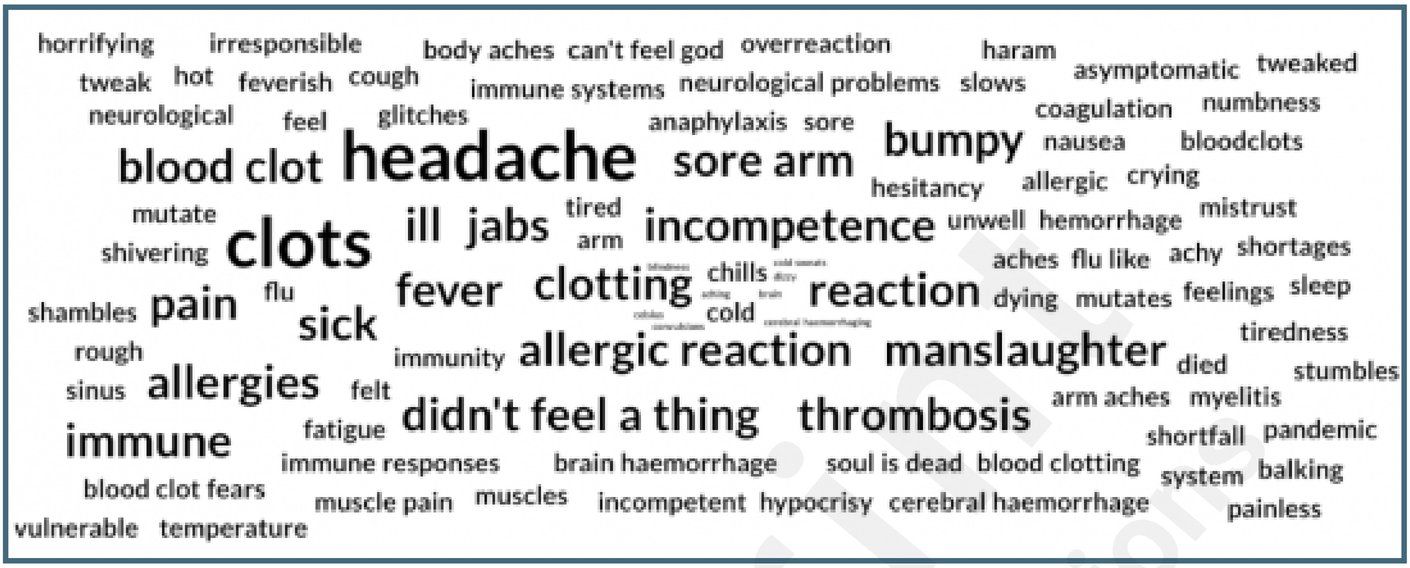
Supplementary Files

Figures

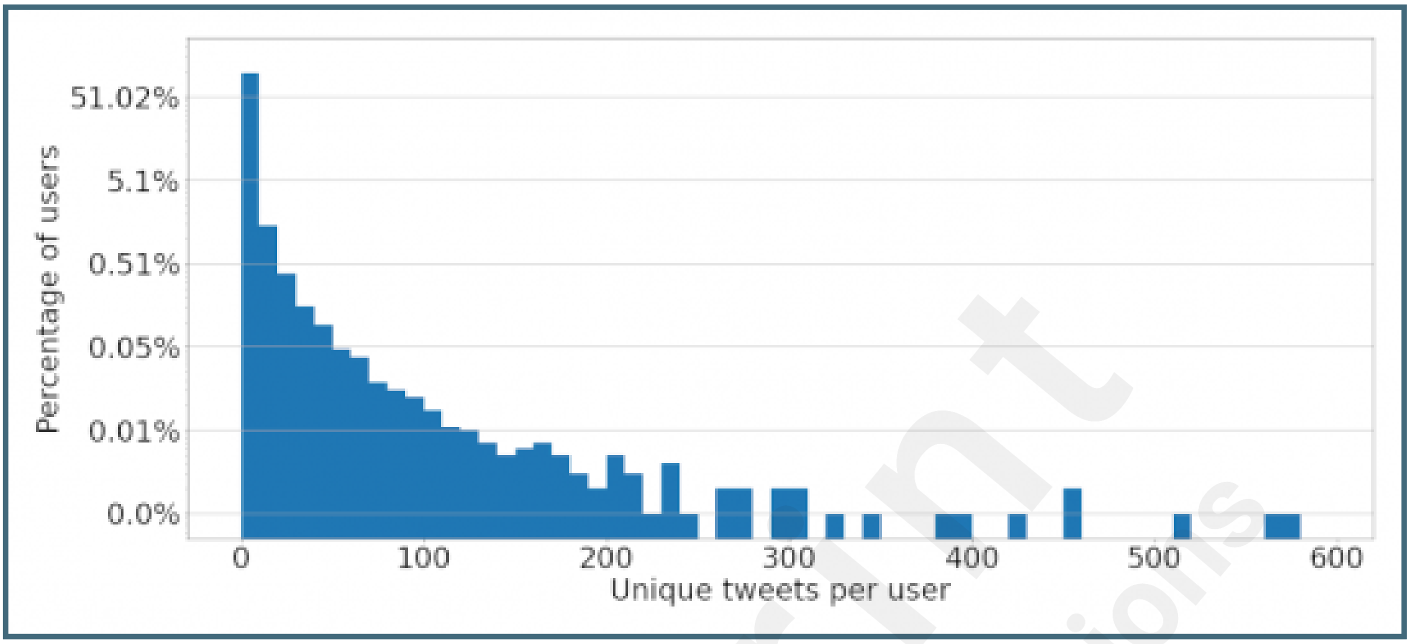
Schema of the full system architecture used to analyze the information displayed on the web portal.



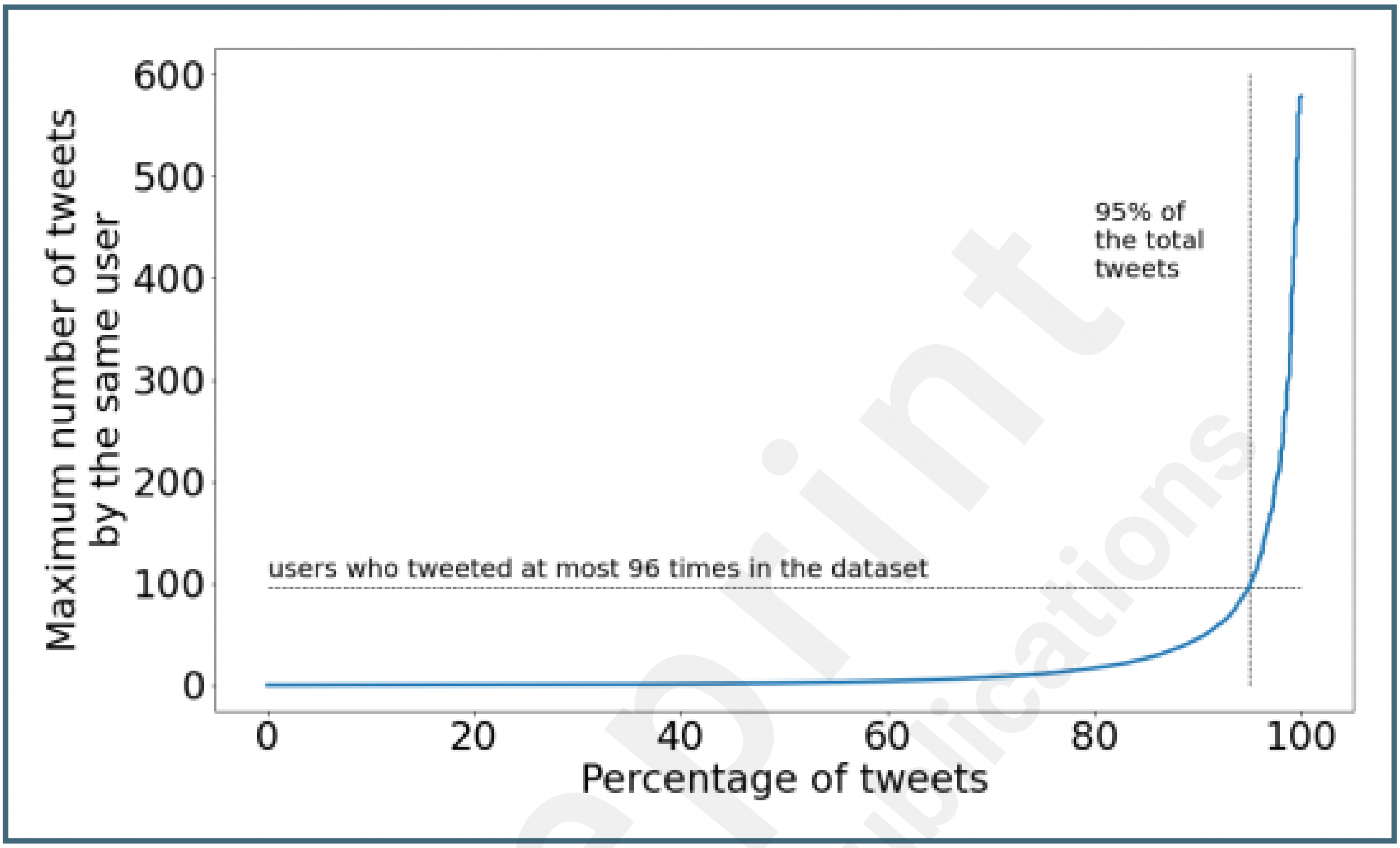
Possible side effects of the AstraZeneca vaccine, as discussed on Twitter. The Word Cloud is generated using our ADE extraction model and displayed on the web portal. The size of the words is proportional to their frequency.



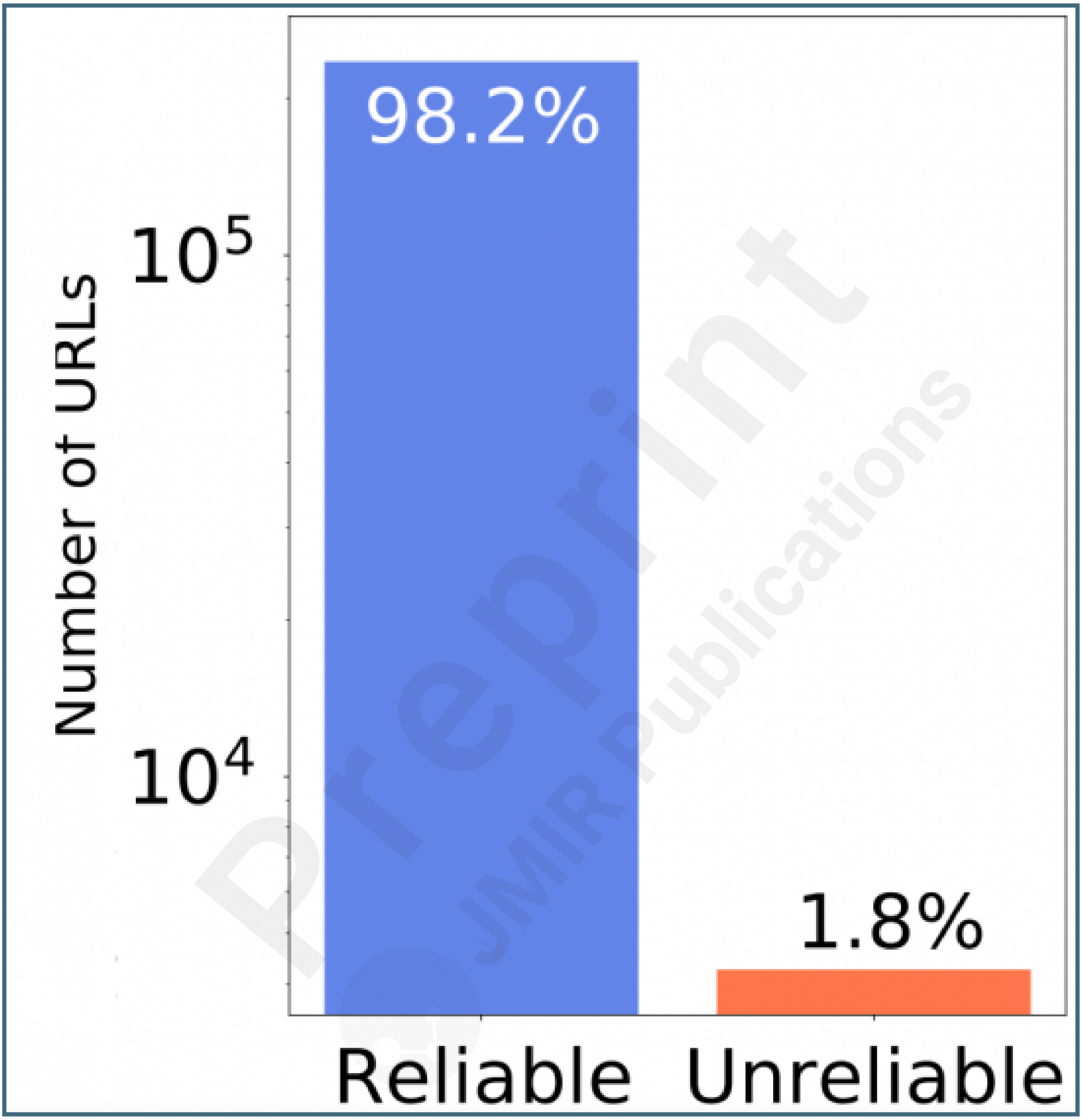
Distribution of users depending on how many times they tweeted (y axes in logarithmic scale).



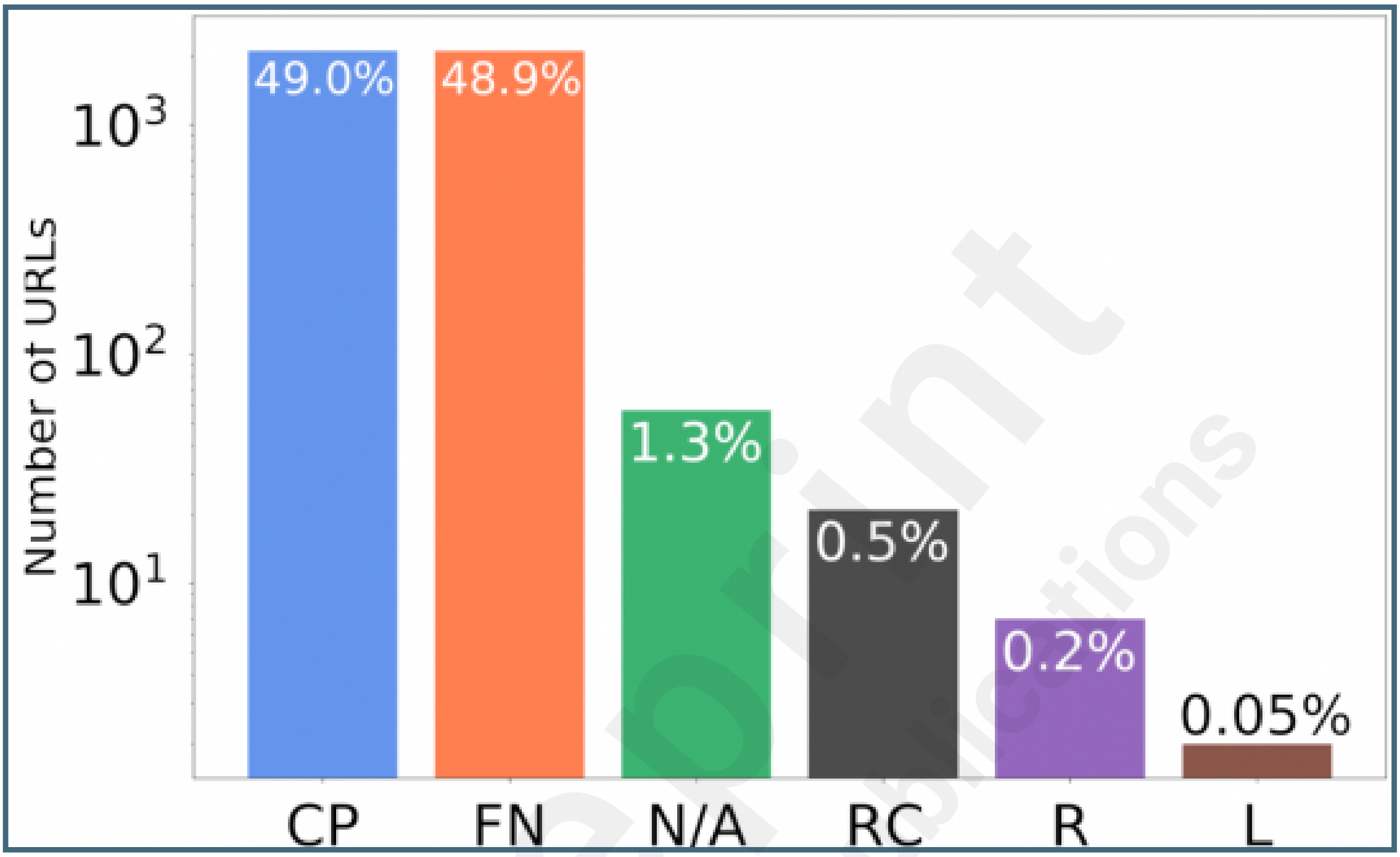
Percentage of tweets produced by a group of users, depending on how many tweets the user produced. 95% of the tweets in the dataset are produced by users who tweeted at most 96 times in the considered timespan.



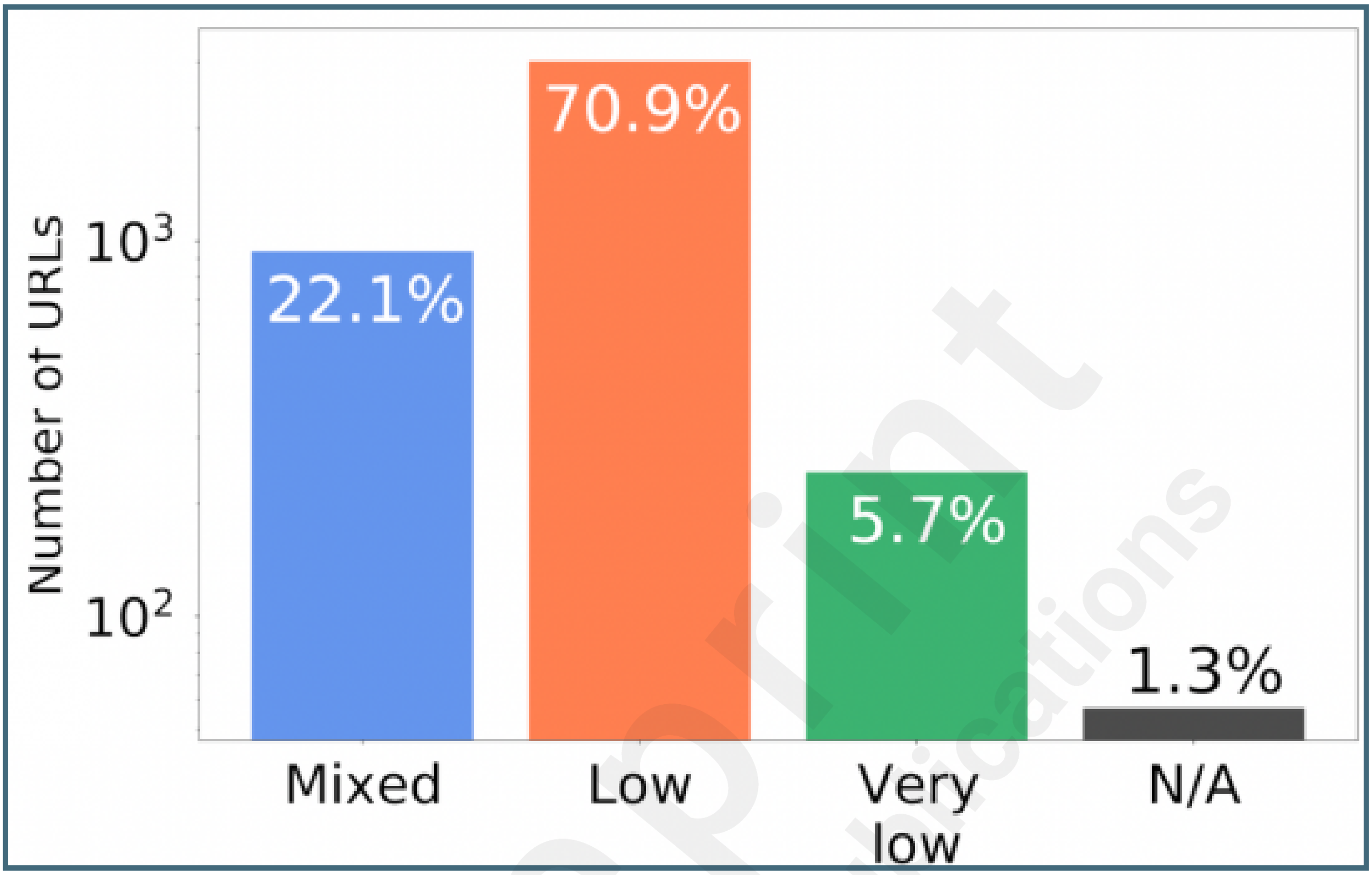
All the plots are in logarithmic scale. Percentage of the Reliable and Unreliable URLs shared.



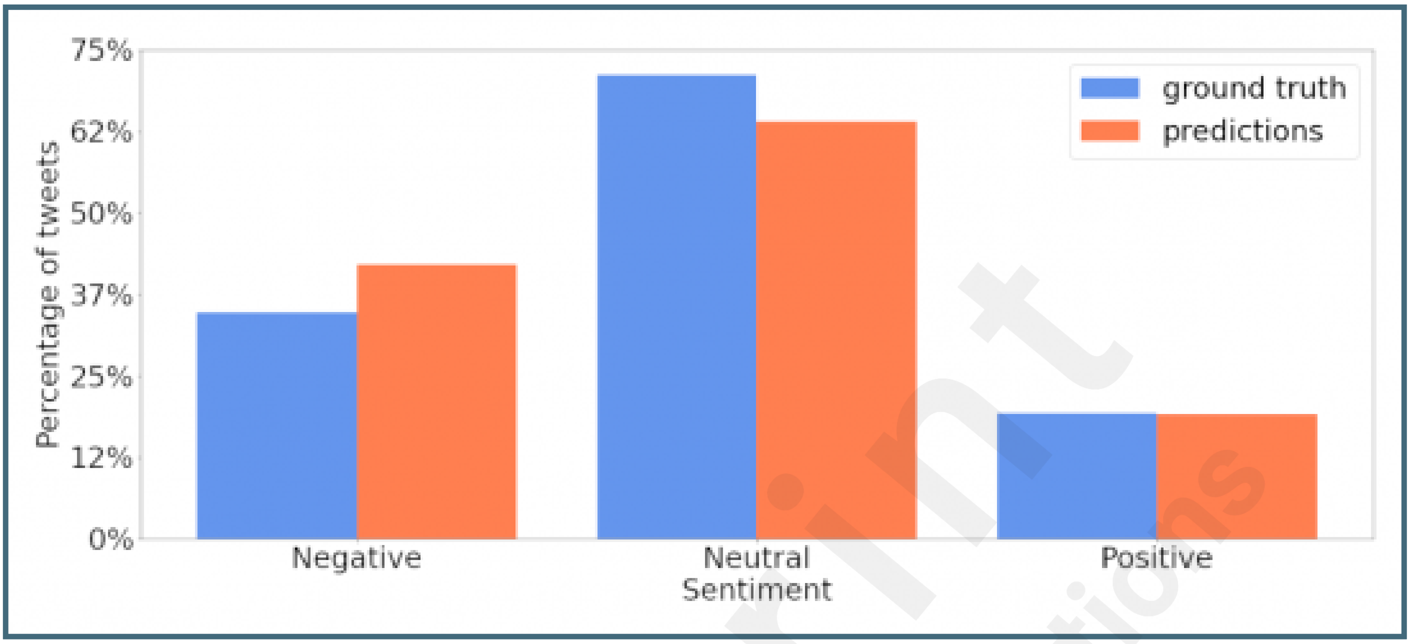
Distribution of MBFC misinformation categories for “Unreliable” URLs: Conspiracy-Pseudoscience (CP), Fake-News (FN), Not Available (N/A), Right-Center bias (RC), Right bias (R) and Left bias (L). y-axis in logarithmic scale.



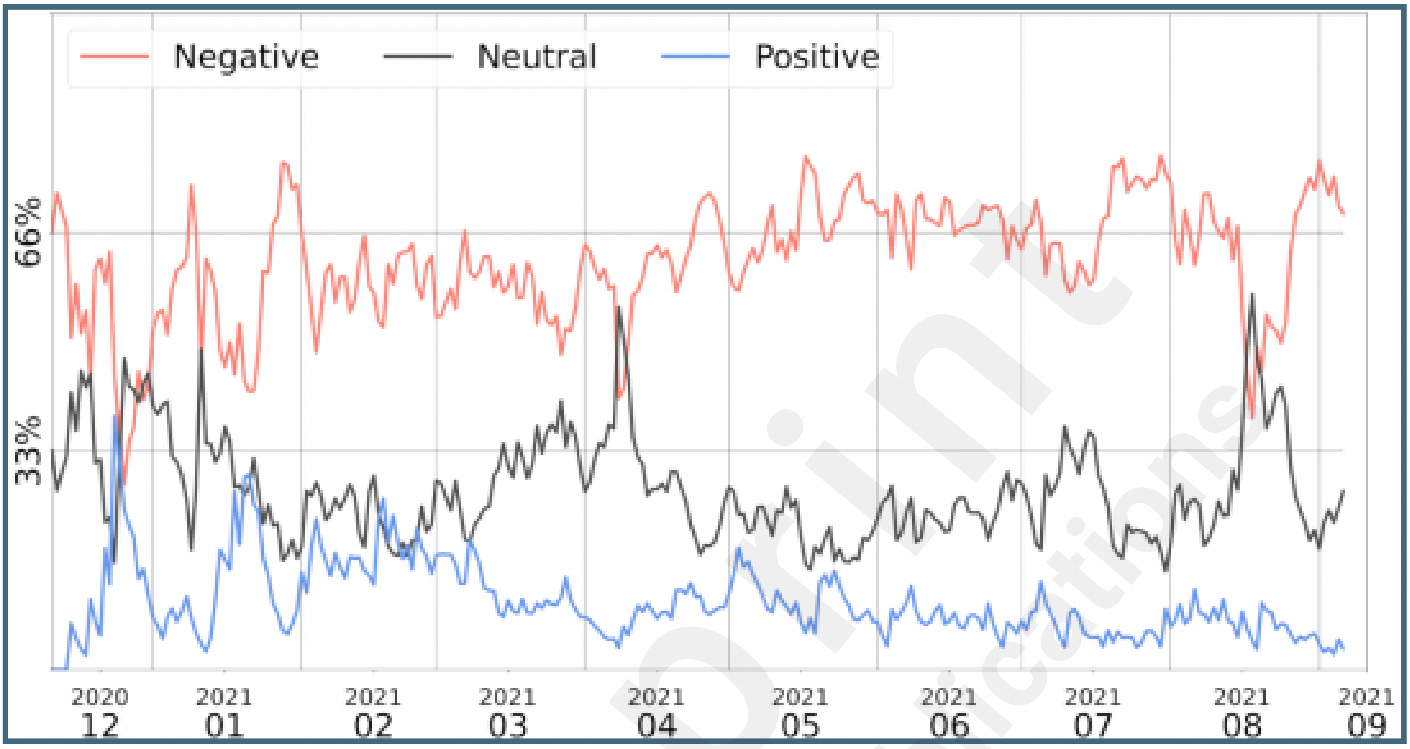
Distribution of MBFC factuality level for “Unreliable” URLs. y-axis in logarithmic scale.



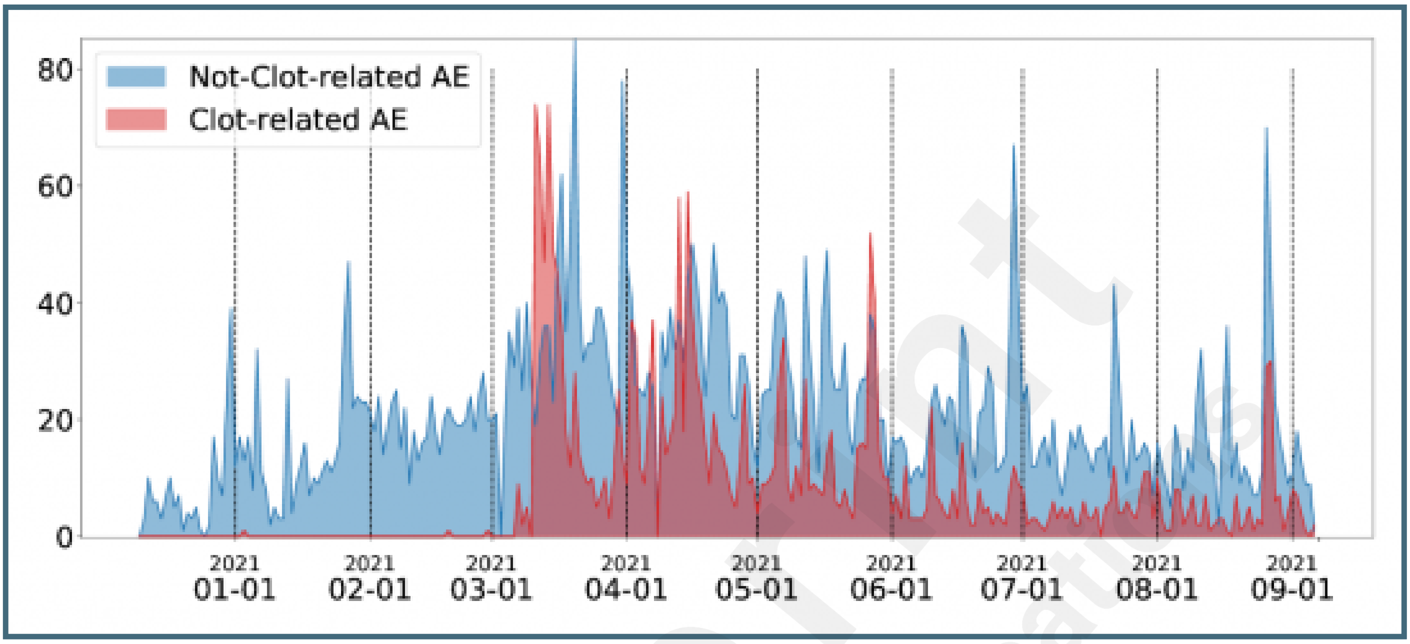
Comparison of the sentiment distributions of the manually-annotated ground truth labels (blue) and the model predictions (orange).



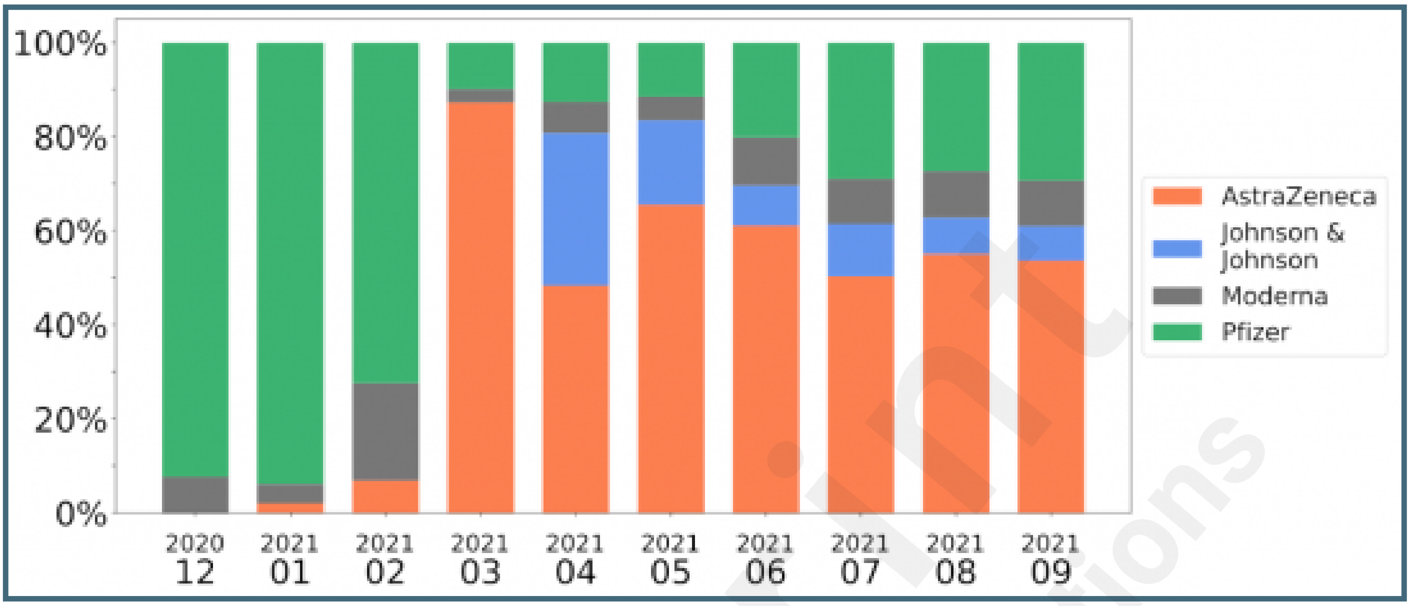
Monthly sentiment distribution in AstraZeneca-related tweets. The y-axis represents the percentage of negative (top, orange), neutral (middle, grey), and positive (bottom, blue) sentiment in the analyzed tweets. It is clear that the prevalent sentiment overall is “negative”, but we can observe spikes of non-negativity in December and January.



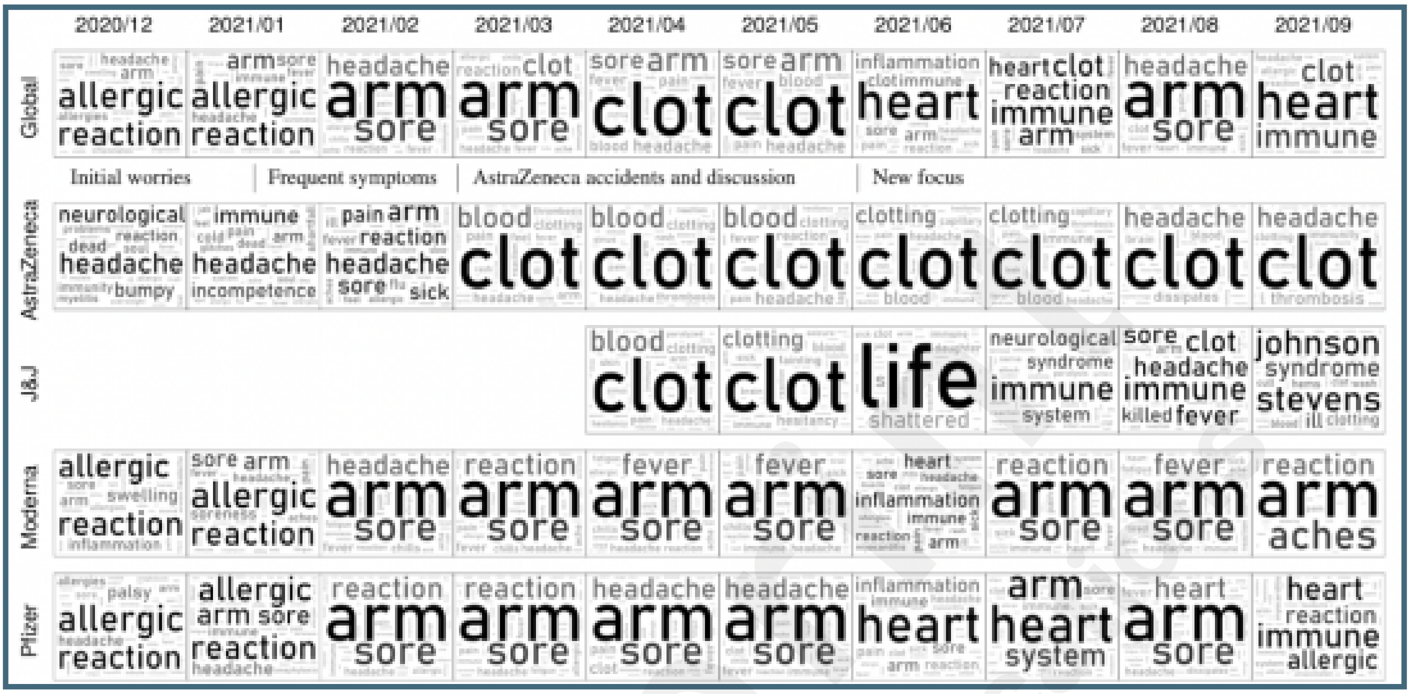
Number of tweets mentioning clot-related and not-clot-related keywords for AstraZeneca (time on the x-axis, number of tweets on the y-axis). The number of tweets mentioning clot-related AEs was initially next to zero, spiked in March 2021 due to media coverage but has been gradually diminishing ever since. Tweets mentioning non-clot-related AEs show a more stable trend over time.



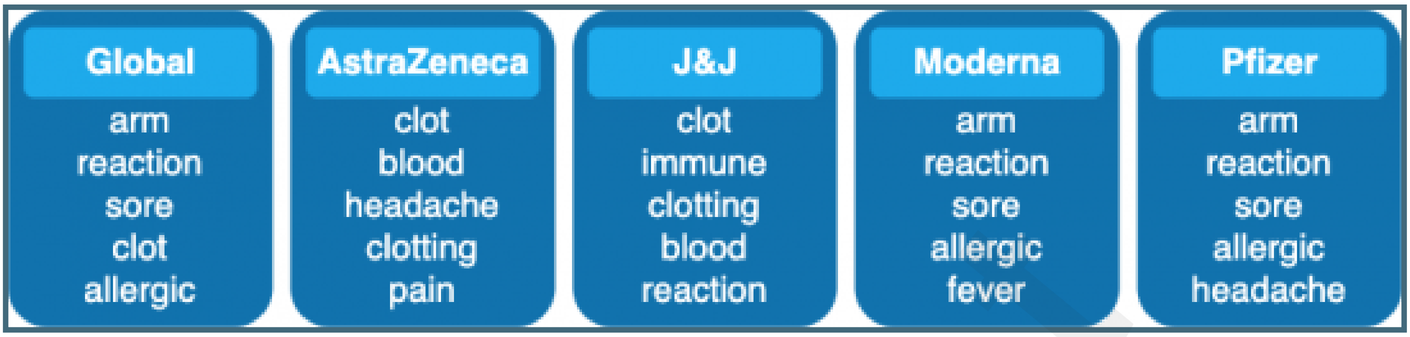
Monthly distribution of vaccine names mentioned in tweets with clot-related keywords (time on the x-axis, percentage of tweets on the y-axis). Most of the tweets were discussing clot-related AEs connected to the Pfizer vaccine before March 2021, when the focus suddenly shifted to AstraZeneca.



Evolution of the global Word Cloud (top row, all vaccines included) and the specific Word Cloud of the following vaccines: AstraZeneca, Johnson&Johnson, Moderna, Pfizer. The suspected ADEs are extracted using our model.



Top-5 most frequently mentioned terms globally and for the following vaccines: AstraZeneca, Johnson&Johnson, Moderna, Pfizer. This takes into consideration all the collected tweets from December 15, 2020 to the beginning of September 2021.



TOC/Feature image for homepages

Schema of the full system architecture used to analyze the information displayed on the web portal.