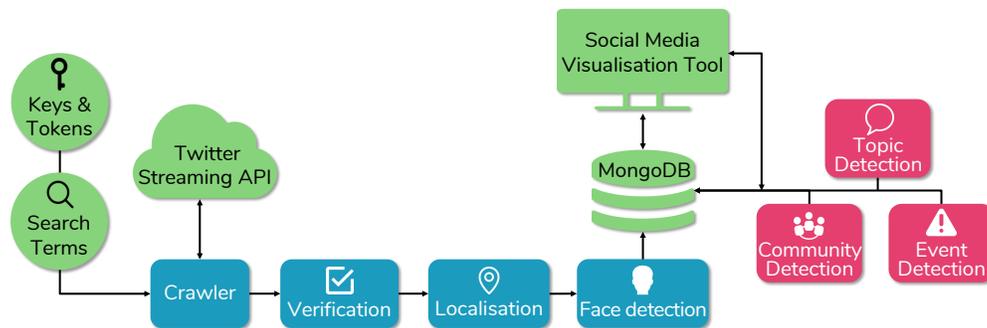

This is a draft version of the article.

The final version will be available at:
<https://www.journals.elsevier.com/online-social-networks-and-media>

Graphical Abstract

A social media analytics platform visualising the spread of COVID-19 in Italy via exploitation of automatically geotagged tweets

Stelios Andreadis, Gerasimos Antzoulatos, Thanassis Mavropoulos, Panagiotis Giannakeris, Grigoris Tzionis, Nick Pantelidis, Konstantinos Ioannidis, Anastasios Karakostas, Ilias Gialampoukidis, Stefanos Vrochidis, Ioannis Kompatsiaris



Highlights

A social media analytics platform visualising the spread of COVID-19 in Italy via exploitation of automatically geotagged tweets

Stelios Andreadis, Gerasimos Antzoulatos, Thanassis Mavropoulos, Panagiotis Giannakeris, Grigoris Tzionis, Nick Pantelidis, Konstantinos Ioannidis, Anastasios Karakostas, Ilias Gialampoukidis, Stefanos Vrochidis, Ioannis Kompatsiaris

- Real-time collection of tweets about the COVID-19 pandemic in highly affected Italy
- Automatic geotagging of tweets and detection of showing faces in their visual content
- Further analysis of tweets to detect trending topics, user communities, and events
- An online platform that visualises the analysed tweets in multi-level aspects
- An interactive map and a visual analytics dashboard to monitor the pandemic crisis

A social media analytics platform visualising the spread of COVID-19 in Italy via exploitation of automatically geotagged tweets

Stelios Andreadis, Gerasimos Antzoulatos, Thanassis Mavropoulos,
Panagiotis Giannakeris, Grigoris Tzionis, Nick Pantelidis, Konstantinos
Ioannidis, Anastasios Karakostas, Ilias Gialampoukidis, Stefanos Vrochidis,
Ioannis Kompatsiaris

Thessaloniki, Greece

Abstract

Social media play an important role in the daily life of people around the globe and users have emerged as an active part of news distribution as well as production. The threatening pandemic of COVID-19 has been the lead subject in online discussions and posts, resulting to large amounts of related social media data, which can be utilised to reinforce the crisis management in several ways. Towards this direction, we propose a novel framework to collect, analyse, and visualise Twitter posts, which has been tailored to specifically monitor the virus spread in severely affected Italy. We present and evaluate a deep learning localisation technique that geotags posts based on the locations mentioned in their text, a face detection algorithm to estimate the number of people appearing in posted images, and a community detection approach to identify communities of Twitter users. Moreover, we propose further analysis of the collected posts to predict their reliability and to detect trending topics and events. Finally, we demonstrate an online platform that comprises an interactive map to display and filter analysed posts, utilising the outcome of

Email addresses: andreadisst@iti.gr (Stelios Andreadis), gantzoulatos@iti.gr (Gerasimos Antzoulatos), mavrathan@iti.gr (Thanassis Mavropoulos), giannakeris@iti.gr (Panagiotis Giannakeris), gtzionis@iti.gr (Grigoris Tzionis), pantelidisnikos@iti.gr (Nick Pantelidis), kioannid@iti.gr (Konstantinos Ioannidis), akarakos@iti.gr (Anastasios Karakostas), heliassgj@iti.gr (Ilias Gialampoukidis), stefanos@iti.gr (Stefanos Vrochidis), ikom@iti.gr (Ioannis Kompatsiaris)

the localisation technique, and a visual analytics dashboard that visualises the results of the topic, community, and event detection methodologies.

Keywords: Social media data, Localisation, Online platform, Knowledge extraction, COVID-19

1. Introduction

Social media have evolved to an integral part of modern society, penetrating the daily life of people around the world. Individuals, from citizens to public figures, as well as organizations of any domain use social media platforms to communicate. The social media posts can share personal details or opinions, inform about public news, promote brands or art, et cetera.

This vast amount of online information available makes social media a reflection of the real world. Major events and trending topics are expressed on social media [1], since individuals now actively participate in the production and diffusion of news [2]. Especially in cases of crisis, social media activity is increased from the early stages till the end of an incident, so as to seek or disseminate emergency information [3].

Analytics in social media data have been proven prominent in crisis management, because they can determine areas in danger, verify events, produce summarized reports, and predict future needs of victims and volunteers [4], greatly contributing to situational awareness and disaster response [5]. When the crisis is a disease outbreak or a virus pandemic, social media can be used to detect the outbreak [6], model the spread of the infection [7], and track the disease [6, 8].

Motivated by the ongoing global COVID-19 pandemic, which was identified in December 2019 and has caused hundreds of thousands deaths, we have designed and developed a framework to collect, analyse, and visualise social media posts that are related to Coronavirus and concern the region of Italy, one of the most deeply affected countries. The main objective is to address the research questions concerning the efficient tailor and consolidation of multiple technologies fostering the multi-level knowledge representation in a well-orchestrated framework. Towards this aim, state-of-the-art technologies are combined to analyse the textual and visual content of the collected social media posts, generating knowledge that visualises into an interactive platform. Specifically, the extraction of spatial information from posts to geo-reference them is another research question that this work deals

with, since it is an essential step to map the posts. The ultimate aim is to provide Italian civil protection authorities, crisis managers, decision/policy makers, and other stakeholders from the health sector with a tool that enables them to monitor relevant social media data, assisting in this way the disease surveillance and the outbreak management.

The social media platform that we have selected to focus on is Twitter, considering its high popularity (330 million active users by the end of 2019¹) and the provided free API for streaming real-time posts, i.e. tweets. Based on keyword search, we constantly collect newly published tweets and analyse them to verify the information they carry, geotag them by detecting locations mentioned in their text, and estimate the number of faces appearing in their images. In addition, we offer techniques that are able to dynamically analyse the Twitter data in order to obtain high-level knowledge, which includes the identification of trending topics, the discovery of user communities, and the detection of events. Finally, we have implemented an online platform that (a) visualises the analysed tweets as pin-points on an interactive map, based on the automatic localisation, offering clustering and filtering capabilities, and (b) provides visual analytics regarding the detected topics, communities, and events.

The main contribution of this paper is a novel framework that consists of a set of state-of-the-art components that analyse data mined from Twitter and a user interface that shows how all these components together support the monitoring of social media activity relevant to COVID-19 crisis in Italy. This contribution can also be broken down to the following points:

- We perform a real-time, keyword-based collection of English and Italian tweets about the COVID-19 pandemic in Italy, in contrast to other works that focus on U.S.A or exclusively on English posts.
- We present and evaluate a deep learning-based localisation technique to automatically geotag tweets, based on locations mentioned in their text instead of the limited geo-information provided by Twitter.
- We apply additional analysis on the textual and visual content of tweets, in order to identify places at risk and cases of overcrowding,

¹<https://www.statista.com/statistics/282087/number-of-monthly-active-twitter-users/>

detect trending topics and influential accounts in user communities, and discover events before or during virus outbreaks.

- We demonstrate an online platform that visualises the collected and analysed tweets in multi-level aspects. An interactive map illustrates tweets, using the outcomes of our localisation technique and face detection, along with clustering and filtering capabilities. A visual analytics dashboard encapsulates the results of the tweets analysis and assists end-users to gain high-level valuable knowledge from the identification of trending topics, events and formed user communities.

The remainder of the paper is structured as follows. Section 2 provides an overview of state-of-the-art works that relate to the exploitation of social media for monitoring the COVID-19 pandemic. Section 3 describes the collection and analysis of social media data, while the developed online platform is presented in Section 4, along with usage scenarios to illustrate its capabilities. Section 5 includes a quantitative evaluation of the localisation, face detection, and community detection approaches, and Section 6 concludes the paper.

2. Related work

The unprecedented pandemic of Coronavirus disease 2019 has dominated the online conversations on social media and the rich content that has been produced publicly is exploited by the research community in several ways.

Analysing the social media data that are posted about COVID-19 can generate further knowledge about the shared information as well as the users behaviour, which can then help policy makers and health care organisations assess the needs of people and address them properly. [9] analysed the engagement of users in relation to the virus by retrieving the activity (likes, comments, shares, etc.), the trending topics and the fluctuation of published content on different social media platforms. [10] created time series of tweets in multiple languages to find the most topically and culturally important words, while [11] focused on tweets posted by world leaders, showing that the majority of posts were informative rather than supportive or political. [12] investigated the public’s attention to COVID-19-related events at the beginning of the pandemic and divided the most discussed topics in five categories: (a) impact, (b) prevention, (c) source of the virus, (d) medical services, and

(e) suspected cases. Similarly, the analysis in [13] grouped the popular topics in four categories: (a) origin of the virus, (b) its sources, (c) impact, and (d) prevention. [14] compared the top topics between March 2020 and January 2021 to show how discussions shifted, while [15] applied topic modeling to monitor topics of concern over time and across areas with socioeconomic disparities. Focusing on user behaviour, [16] performed network analysis on Korean Twitter data and showed that information was spreading faster in networks that referred to “Coronavirus”. Our work involves knowledge extraction techniques similar to the aforementioned, such as the detection of trending topics and user communities, but is also extended to recognise locations mentioned in the tweet texts and the number of people (faces) in the tweet images.

Unfortunately, the increase of shared information about COVID-19 has led as well to the spread of unrestrained misinformation and conspiracy theories over traditional and social media. Considering that a large number of people rely on social media platforms for news and that information from both reliable and questionable sources present similar spreading patterns [9], the identification of misinformation has emerged as a critical task. [17] proposed a COVID-19 fake-news detection model with robust loss and influence-based cleansing, while for the same problem [18] proposed a cross-stitch based semi-supervised end-to-end neural attention model. [19] have designed a dashboard to track misinformation on Twitter based on the credibility of the news sources of the shared links. [20] analysed the magnitude of false information spread on Twitter to discover that one out of four tweets included misreports about public health, while [21] identified the main types, sources, and claims of Coronavirus misinformation and found that it originates from ordinary people, usually with allegations against authorities. Moreover, [22] provided evidence that Twitter bots are used to promote political conspiracies in the USA and [23] found that low-credibility information spreads mostly via retweets and, although bots are involved, most of the content is generated by humans. On a positive note, [24] found that fake news are discussed at a lower volume than other conversations about the virus and [25] revealed through social network analysis that, despite the fact that a certain conspiracy was popular, only a limited number of users actually believed it. In order to tackle the problem of online misinformation, we adopt an automatic classification model that takes into account characteristics of a tweet and its author, and estimates whether it is real or fake.

The ability to predict the development of a disease outbreak as early and

reliably as possible is invaluable to the prevention of its spread. [26] demonstrated that the peak interest for Coronavirus-related keywords in search engines and social media data was 10-14 days before the incidence peak of the virus published by official authorities, and [27] proposed a fuzzy rule-based evolutionary model to run on crawled tweets in order to timely detect outbreaks of the virus. Our event detection methodology is able to produce early alerts when there is an increase on the number of tweets we collect.

The creation and publication of social media archives greatly facilitate the research community in identifying misinformation, measuring the public sentiment, and predicting outbreaks, by providing a dataset for training and testing of machine learning and information retrieval methodologies. [10] and [28] published multilingual Twitter datasets about COVID-19, while [29] introduced a dataset in Arabic. Furthermore, [30] focused on tweets that were posted during the early stages of the virus spread and [31] as well as [32] offered large-scale curated datasets of over 150 and 198 million tweets respectively. Other datasets are CTF [18] with more than 45 thousand tweets labelled genuine or fake and COVIDSenti [33] with 90 thousand tweets labelled into sentiment classes. Another online dataset² contains the geographical distribution of tweets, but it was solely based on the geo-coordinates provided by Twitter. Alternatively, [34] provided a publicly available knowledge base of more than 8 million tweets, exposed using established RDF/S vocabularies and enhanced with extracted sentiments. At the moment of writing, we have analysed more than 40,000 Twitter posts that refer specifically to the virus spread in Italy and enhanced them with geographic information that derives from our automatic localisation technique. The collected tweets can all be displayed and filtered in our Social Media Visualisation Tool.

Other works concern the visualisation of crawled posts towards a user-friendly monitoring of COVID-19 over social media. To the best of our knowledge, three projects are relied basically on Twitter data and display them in an interactive environment, thus we briefly discuss them in the following paragraphs.

The *CoronaVirusTwitterMap*³ tailors interactive visual analytics technologies for processing and illustrating a large amount of COVID-19-related

²<https://data.humdata.org/dataset/covid-19-twitter-data-geographic-distribution>

³<https://coronavirustwittermap.ics.uci.edu/>

Twitter data and case numbers. It has been developed by the computer scientists team of the University of California, Irvine (UCI). The spatial and temporal distribution of tweets related to the pandemic is illustrated through an interactive map that is mainly focused on Twitter activity in the U.S.A. and U.K. currently. The tweets are collected through TwitterAPI since January 2020 and stored in real-time to a database, while the case numbers are acquired from the 1Point3Acres.com⁴ project [35]. The aim of this project is to enable users to view the increasing social media activity as the contagion spreads, without any further back-end analysis on the tweets' content being realised.

The *OmniSci analytics platform*⁵ originates as an outcome from MIT research and experience. It delivers an interactive way to explore and visualise large datasets in real-time by dealing with analytical questions in short time [36]. Furthermore, the *OmniSci Tweet Map*⁶ is an interactive dashboard that enables users to define their own hashtags and visualise the resulted tweets on a list or over the world providing geographical mapping information. Furthermore, the *OmniSci interactive demo for COVID-19*⁷ provides a visual overview of confirmed cases and spread in the U.S.A. by exploiting the data collected through the *COVID Tracking Project*⁸. However, this interactive demo is not based on Twitter information and the OmniSci Tweet Map is not dedicated to Coronavirus.

Finally, the *COVID-19 Real Time Tweet Map*⁹ is a map-based Web application that shows which locations tweet the most about Coronavirus, along with a live stream of new tweets and an estimation of the sentiment they carry. Nonetheless, users are able to view exclusively tweets that have been posted after visiting the application, and the analytics are limited to the most common words used in the tweets.

In contrast to the aforementioned works, we have developed a map-based visualisation of tweets collected in real-time, which is dedicated to monitoring Coronavirus in Italy and utilises the automatically detected locations rather than Twitter's insufficient geo-information, and a visual analytics dashboard

⁴<https://coronavirus.1point3acres.com/>

⁵<https://www.omnisci.com/platform>

⁶<https://www.omnisci.com/demos/tweetmap>

⁷<https://www.omnisci.com/demos/covid-19>

⁸<https://covidtracking.com/>

⁹<https://www.1point21interactive.com/maps/covid-map/>

that displays the most discussed topics, influential users and alerting incidents.

3. The proposed framework

The architecture of the proposed framework that collects, analyses and visualises COVID-19-related tweets is illustrated in Fig. 1. The core component of the framework is the Crawler (Section 3.1), which establishes a connection to the *Twitter Streaming API*¹⁰ and continuously receives new tweets that satisfy predefined search criteria. For every collected post a three-step analysis is performed to verify the tweet by estimating the reliability of the incoming information (Section 3.2), to localise the tweet by detecting the locations mentioned in its text and associating them to coordinates (Section 3.3), and to find the number of people appearing in its image by detecting faces (Section 3.4). The outcomes of the analysis are added as complementary attributes to the original tweet, which is finally stored to a secure database. Supplementary analysis techniques, implemented as independent APIs, connect to the database and use tweets as their input so as to track topics (Section 3.5), discover user communities (Section 3.6), and identify events (Section 3.7). The *Social Media Visualisation Tool* (Section 4) searches directly in the database to fetch the tweets to be displayed or calls the analysis APIs to retrieve and visualise higher knowledge.

3.1. Collection of social media data

Amongst a multitude of available Twitter API endpoints, the Streaming API can be considered the most suitable for our scope, since it allows access to Twitter’s global stream of data and retrieves public tweets almost at the instant they are posted. In order to specify what tweets should be retrieved from the whole stream, the API offers three filtering options: user accounts to follow, keywords to appear in the messages and bounding boxes to track.

In our case, we have selected to search by keywords, and specifically the popular Twitter hashtags *#COVID19Italy* and *#COVID19Italia*. We have intentionally avoided using search by bounding boxes, because we are more interested in tweets that refer to the examined area, i.e. Italy, rather than in the coordinates they have been posted from.

¹⁰<https://developer.twitter.com/en/docs/tweets/filter-realtime/api-reference/post-statuses-filter>

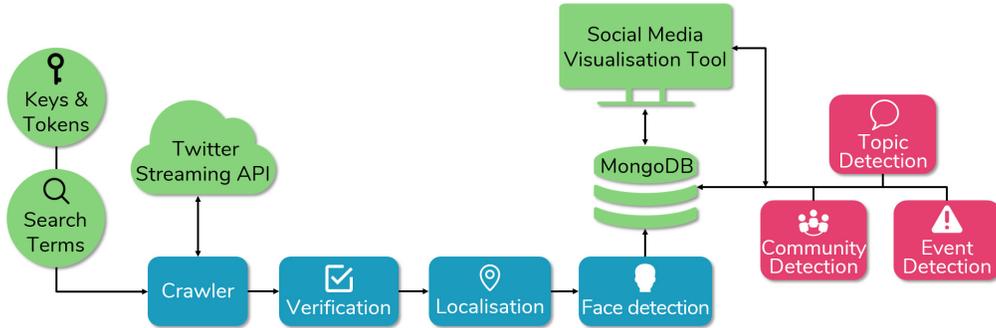


Figure 1: The proposed framework of collection, analysis, and visualisation of COVID19-related tweets

Each tweet that is consumed from the API is analysed by three different methodologies in order to (a) verify the credibility of the tweet, (b) geotag the post based on mentioned locations, and (c) find the number of faces appearing in its image. The outcomes of these analyses are stored along with the tweet and its metadata.

After collecting tweets with the hashtags “#COVID19Italy” and “#COVID19Italia” for circa three months (from May 7, 2020 to July 22, 2020), the status of the dataset can be seen in Table 1. More than 40 thousand tweets have been collected, with the majority being in Italian and a small percentage in English. The size of the collection might seem limited in comparison to other datasets of related works, but the number is reasonable taking into account that our use case is focused only on Italy and for a short period of time. Over half of the posts are retweets and one out of two also contains an image. It should be highlighted that only 0.4% of the posts include their original geographical information (derived from Twitter), while our localisation technique has achieved to detect locations in the text of circa 14,000 tweets.

Total	English	Italian	Retweets	With image	Twitter location	Detected location
43,631	3,028 (7%)	40,603 (93%)	26,171 (60%)	20,144 (46%)	162 (0.4%)	14,423 (33%)

Table 1: Status of the collection after three months of crawling

3.2. Verification

Regarding the credibility, an automatic verification method [37] [38] is applied on the collected tweets to handle fake news and misinformation. This solution relies on two independent classification models built on the training data using two different sets of features: tweet-based and user-based. Tweet-based features can be (a) text-based (e.g., number of uppercase characters), (b) language-specific (e.g., number of detected slang words), (c) twitter-specific (e.g., number of retweets, number of hashtags), and (d) link-based (e.g., existence of external URLs). On the other hand, user-based features can be (a) user-specific (e.g., number of user’s followers) and (b) link-based (e.g., existence of a ULR in the Twitter profile description). Following the feature extraction, model bagging is used to produce more reliable predictions based on classifiers from each feature set. The classification algorithms are Logistic Regression and Random Forests of 100 trees. In addition, at prediction time, an agreement-based retraining strategy is applied, which combines the outputs of the two bags of models in a semi-supervised learning manner. The final output is a binary classification of the tweet as “real” or “fake”, together with a confidence score (0-1), which is finally transformed to a reliability score as follows:

$$reliability = \begin{cases} x, & \text{if } prediction = real \text{ with confidence } x \\ 1 - x, & \text{if } prediction = fake \text{ with confidence } x \end{cases}$$

For example, the method predicts that the tweet in Fig. 2 is fake with confidence 0.69, due to the extensive usage of emojis, the external URL and the significantly low number of followers of the account, and thus the reliability score is 0.31. On the other hand, the prediction for the tweet in Fig. 3 is that it is real with confidence 0.75, since its text is clean, the attached link is safe and the account has many followers; respectively, the reliability score is 0.75.

3.3. Localisation

This section presents a deep learning-based approach that retrieves location- and organisation-type mentions in the short texts of tweets in the examined languages, i.e. English and Italian. A Named Entity Recognition (NER) framework handles the task of retrieving the candidate entities, which are then used to populate a territory map with the respective tweet locations.

Komplette Video in Bio YouTube



#Coronavirus #Sicilia #Siacca #Agrigento
#CoronavirusItalia #Covid19 #Covid19Italia
#eccellenzesiciliane #موثق #كرونا #كورونا #coronavirus
#virus #ويروس #... [instagram.com/p/CBf8VvHIHJN/...](https://www.instagram.com/p/CBf8VvHIHJN/)

5:23 PM · Jun 16, 2020 · Instagram



Figure 2: Example of a tweet predicted as fake with confidence 0.69

Italy opens borders to European tourists as coronavirus lockdown lifts
#COVID19 #COVID2019 #COVID19italia #COVID—19
#coronavirus #coronavirusitalia #coronavirusitaly



2:26 PM · Jun 3, 2020 · Twitter Web App



Figure 3: Example of a tweet predicted as real with confidence 0.75

Specifically, the methodology exploits a variety of Bidirectional Long Short-Term Memory (biLSTM)-based models which are evaluated with established datasets for each treated language.

One of the features that differentiates the presented method from other approaches, other than the English/Italian focus, resides in the exploitation

of the “organisation” class in addition to the more traditional “location/geopolitical entity” ones. By exploiting organisation names found in Twitter text, a tweet’s respective location can be retrieved by searching the aforementioned names’ geographic coordinates using relevant online mapping services. Naturally, probable location names found in the same tweet are also used as criteria to pinpoint the desired location, since sometimes only considering an organisation’s name is inconclusive (as is usually the case with conglomerates consisting of many subsidiaries).

Modern state-of-the-art NER models have been tested in both supported languages. An analysis of the neural networks’ architectures and main characteristics follows to illustrate their most important features. To efficiently represent both the placement and the form of an entity in a candidate sentence, word-level and character-level representations are exploit.

While all of the chosen models rely on biLSTM-based architectures, various implementations have been tested (BiLSTM-CRF[39], BiGRU-CRF[40], BiLSTM-CNNs-CRF[41], BiLSTM-CNN[42]) to determine the best possible outcome. LSTMs are ideal when dealing with sequences of data (like text) because of the way the network channels information via its nodes, leveraging its ability to keep information in memory based on history. Any sequential information is kept in the LSTM’s internal state (i.e. hidden layer) and is updated with each new data via input/output and forget gates. This way the network is capable of predicting the output based on long distance dependencies. The bidirectional nature of the LSTM network manifests with two processes, applicable to each lexical unit of a given sentence that each computes a representation of the lexical unit’s left and right context.

To successfully manage the extraction of multilingual location and organisation entities, established datasets are being exploited for the training of the tested models.

Concerning the English language, the CoNLL2003 dataset [43] was used for training and testing of the respective model. The included annotation caters for four named entity variations: location (LOC), organisation (ORG), person (PER) and miscellaneous (MISC). To fully exploit the given tagged information, instead of only focusing on the “location” entities, it was decided to explore the use of the “organisation” ones as well, in a bid to infer possible locations out of organisations’ locales.

The dataset used for training the NER model in the Italian localisation pipeline, was introduced in the respective EVALITA 2009 track [44]. The annotation labels are similar to the ones used in the English dataset, with the

GPE (Geo-Political Entity) entity category replacing the MISC one, found in CoNLL2003, as a main differentiation. The respective annotation format also follows the BIO/IOB2 schema, which was found to be advantageous in the specific implementation, since LSTMs favour its simplicity [45].

The parameters that were used during the training of the Bi-LSTM-CRF models are displayed on Table 2. The word representations used as input in the current model are the publicly available, pre-trained GloVe embeddings [46] for the English language, while for Italian the respective pre-trained fastText [47] embeddings were used. The same settings were applied to both languages.

Training parameters	Value
LSTM size	100
Batch size	20
Dropout rate	0.5
Epochs	25
Optimiser	Adam
LSTM size	100
Character embeddings dimensions	100
Word embeddings dimensions	300

Table 2: Bi-LSTM-CRF parameters used for both languages

After a word has been recognised as a location or an organisation entity, the point as well as the bounding box of the detected location are retrieved utilising the *OpenStreetMap API*¹¹. Usually, OpenStreetMap returns multiple responses for an entity and lists them based on their popularity. We consider the most popular one and retrieve the coordinates of its bounding box and the coordinates of its specific point. In case of multiple recognised locations, we compare the bounding boxes with each other and if one is included in the other, only the more precise is considered. If the bounding boxes do not overlap, all of the points of the detected locations are considered and the tweet is geotagged multiple times, so as to be displayed in all the places of the map it refers to.

¹¹https://wiki.openstreetmap.org/wiki/API_v0.6

3.4. Face detection

Face detection is the process of automatically discovering the locations of all the human faces that appear in an image. This is done typically by finding boxes that tightly enclose the faces in the image. The boxes are described by assigning pixel coordinates for the top, bottom, right and left edge of each box. Early face detection techniques were based on handcrafted features like the Viola-Jones Haar cascades, Histograms of Oriented Gradients, or Local Binary Patterns [48, 49, 50]. Those were quickly abandoned in the era of deep learning and replaced with deep architectures that achieved very high accuracy rates in known benchmarks [51, 52, 53, 54]. In this work, face detection is used to assist to the detection of crowded places. Crowd analysis requires excellent face detection accuracy in difficult cases of crowded scenes that may pose the challenges such as occlusion, illumination and scale. Thus, we adopt the TinyFaces face detector [55] which is specifically designed to be robust to scale.

The TinyFaces model in essence was designed to solve a binary multi-channel heatmap prediction problem, where the heatmap values at a single pixel location represent face scores for boxes with various window sizes centered around the pixel. In order to accomplish this, a Fully Connected Network was designed with multiple scale-specific windows (templates).

In an extensive study by [55], it was proven that using deep CNN features from multiple layers, and thus information from different receptive fields, is important so as to find small faces. This means that the visual context that surrounds the faces inside the candidate boxes play a significant role in determining the existence of a face. Practically, it seems natural even for humans, to require some of that surrounding context to accurately infer whether a real size 25x20 pixel portion of an image resembles a face. In contrast, in order to detect bigger faces (e.g. 300x300), it is enough to get no context at all but only what is provided inside the face portion.

In the same manner, resolution can have different impact on performance for various template scales. It was proven that building templates at the original resolution is not optimal. For small faces, the scale-specific template size is doubled in combination with 2x up-sampling in resolution, improving accuracy by 6.3% compared to standard template size with standard resolution. This translates to the opposite for bigger faces: when templates and resolution were downscaled by a factor of 2, an improvement in accuracy by 5.6% was observed. This is done in order to accommodate the natural overfitting towards medium-sized objects a pre-trained CNN architecture has

been exposed to, as a result of the fact that medium-sized objects are dominant in large scale datasets. The TinyFaces pre-trained model we acquired is fine-tuned on the WIDER FACE dataset [56].

The final output of the face detector is only a single number per image, which refers to the number of faces detected. It is not required to store the images in the system, therefore they are deleted once the analysis is complete along with any extracted features and metadata. Please note that the system is not capable of performing person identification and the processing stops as soon as we detect the faces. Some examples can be seen in Fig. 4, where the face boxes found by the detector are drawn using blue rectangles, but only for visualisation purposes in this paper.

3.5. Trending topics

Topic detection refers to the clustering of textual streams of data into groups of similar content. The work in [57] proposes a combination of density-based clustering with Latent Dirichlet Allocation (LDA) [58]. First, the module estimates the number of clusters (topics) [59] and then the estimation is followed by LDA to assign social media posts to topics.

The estimation of the number of topics is performed in an incremental way on the vector representation of short text using word n -grams for $n = 1, 2$. A set of density levels $\epsilon_t, t = 1, 2, \dots, T$ is randomly generated. Afterwards, for each density level ϵ_t a clustering output $C_{DBSCAN(\epsilon_t)}$ is obtained in multiple processing nodes in a parallel way from a set of mutually orthogonal vectors $C^{(t)}, t = 1, 2, \dots, T$ defined as follows [59]:

$$C^{(t)}[j] := \begin{cases} 0 & \text{if point } j \text{ belongs to a previously extracted} \\ & \text{cluster at density levels } \epsilon_1, \epsilon_2, \dots, \epsilon_{t-1} \\ C_{DBSCAN(\epsilon_t)}[j] & \text{otherwise} \end{cases}$$

where $C^{(1)} = C_{DBSCAN(\epsilon_1)}$. The final set of density-based clustering outputs $C_{DBSCAN(\epsilon_t)}, t = 1, 2, \dots, T$ are combined to get the final estimation of the number of topics $\hat{k} = \text{DBSCAN-Martingale}(\text{minPts}, T)$.

The output of the above methodology is the detected clusters/topics, the tweets each cluster comprises and the most frequent words for each topic.

3.6. User communities

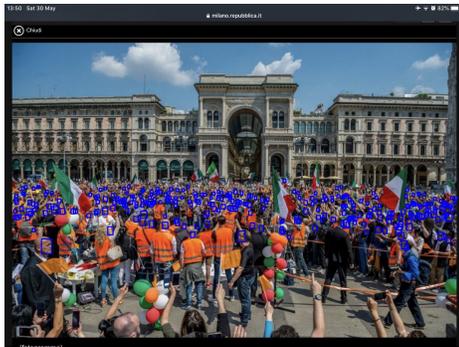
Community detection is the discovery of communities of social media users that are interlinked through their online behaviour, i.e. user accounts



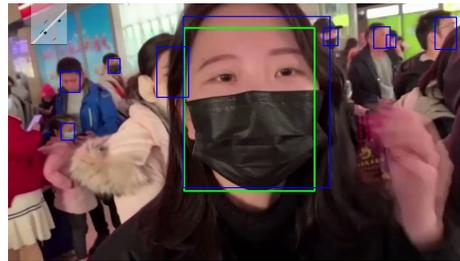
(a)



(b)



(c)



(d)

Figure 4: Faces detected in COVID-19 social media images.

that mention each other on Twitter. This relationship has been selected over other, e.g. following, since it defines a more temporary connection of users and user communities shift continuously based on discussed topics and events. Complementarily, key-players identification is the discovery of the most influential accounts, based on their position in the network graph.

The methodology [60] starts by denoting the social network as $G(N, L)$, where N nodes represent the Twitter user accounts and L the links between them; a link (i, k) means that user n_i mentions or is mentioned by user n_k . The next step is to apply a community detection algorithm in order to divide the network into groups of users who are more densely connected to each other within the group rather than to the rest of the network. After an evaluation of alternative community detection algorithms (see Section 5.3), the fast and scalable Louvain algorithm [61] has been adopted. The final output of the methodology is the number of detected communities and the set of nodes that belong to each community.

To identify the key-players, an entropy-based centrality measure is applied, i.e. the Mapping Entropy Betweenness (MEB) centrality, which considers the betweenness centrality of nodes [62]. Specifically, the betweenness centrality (BC) of node n_k is based on the number of shortest paths $g_{ij}(n_k)$ from node n_i to node n_j that pass through node n_k to the number of all shortest paths g_{ij} from node n_i to node n_j , summed over all pairs of nodes (n_i, n_j) and normalised by its maximum value, i.e. $(N^2 - 3N + 2)/2$:

$$BC_k = \frac{2 \sum_{i < j}^N \frac{g_{ij}(n_k)}{g_{ij}}}{N^2 - 3N + 2}$$

When a node acts as a bridge between many pairs of nodes, its betweenness centrality is relatively high. To that end, MEB centrality focuses on the betweenness centrality of a node, but also considers the betweenness centrality of its first neighbours. MEB is defined as follows:

$$MEB_k = -BC_k \sum_{n_i \in N(n_k)} \log BC_i$$

where the weight assigned to BC_k is the sum of all $-\log BC_i$ over the neighbourhood of node n_k . Ultimately, the nodes with the maximum MEB centrality are considered key-players.

3.7. Events

Event detection aims to identify real-world incidents by discovering anomalies in streams of social media data. In particular, notable differences (outliers) in the fluctuation of the number of tweets that have been collected per day can be associated to events.

To detect these outliers, the method of z-score [63] is applied. Z-score indicates how many standard deviations a data point is from the sample's mean (assuming a Gaussian distribution) and, when it exceeds a constant threshold, the data point is considered an outlier. The formula to calculate the z-score is the following:

$$z = \frac{x - \mu}{\sigma}$$

where x is the value of the examined point, μ is the mean of a sample x_1, x_2, \dots, x_n , i.e.

$$\mu = \frac{1}{n} \left(\sum_{i=1}^n x_i \right)$$

and σ is the standard deviation of the sample, i.e.

$$\sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \mu)^2}$$

To characterize a data point as an outlier, its z-score has to be above a specified threshold t (bibliography suggests a value between 2.5 and 3.5):

if $z > t$ then x is outlier

In order to apply the outlier detection methodology to our problem, i.e. the detection of events through social media data, we calculate the z-score of each date. The data points refer to the number of collected tweets per date and each point is compared to the other points of the last 30 days, while the threshold has been set to 3. When an event is detected, additional analysis is performed, e.g. the extraction of the most mentioned words, so as to gain an insight on what that date's incident really is about.

4. Online platform

The *Social Media Visualisation Tool*¹² encapsulates in a multi-level framework the aforementioned components for social media analysis enhancing the efficient management of the pandemic crisis. Specifically, it consists of two parts, namely the *Interactive Map* and the *Visual Analytics Dashboard*. The former is an interactive, user-friendly interface that facilitates the navigation through a geographical map and exploits the zoom and search functionalities, in order to present valuable and actionable information that relies on localisation and face detection techniques upon the individual tweets. The latter illustrates the social media analysis outcomes and the high level extracted knowledge by employing comprehensive visualisations.

The presented Web-based and open source application has been inspired by the elemental ideas and aspects of various EU funded projects, such as H2020 beAWARE¹³, H2020 EOPEN¹⁴ and H2020 aqua3S¹⁵. A presentation of the online platform and a description of usage scenarios are carried out in the following subsections.

4.1. Interactive Map

The cornerstone of the Social Media Visualisation Tool is the interactive geographical map (Fig. 5) in which the incoming incidents, i.e. newly published Twitter posts, are visualised as single blue pin-points or grouped formulating bigger clusters depending on their vicinity. For example, in Fig. 6 some single incidents as well as clusters of posts are presented. Users are able to navigate to each one of the clusters and explore its members' content. The color of a cluster indicates its cardinality (number of members that belong to it). For each single incident, its description, Twitter identification number (ID) and any attached images are displayed in a pop-up that shows when a pin-point is clicked.

The platform processes tweets that are written in English (EN) as well as in the Italian (IT) language, as presented in the "Details" section (Fig. 7). The section also provides quick information concerning the number of analysed tweets in the current month ("Tweets this month"), in the previous

¹²Available online at <http://beaware-server.mklab.it:3000/>

¹³<https://beaware-project.eu/>

¹⁴<https://eopen-project.eu/>

¹⁵<https://aqua3s.eu/>



Figure 5: An overview of the Social Media Visualisation Tool

(completed) month (“Tweets previous month”), and the total amount of processed tweets since the beginning of crawling (“Total Tweets”). It also informs about the number of analysed tweets per each language.

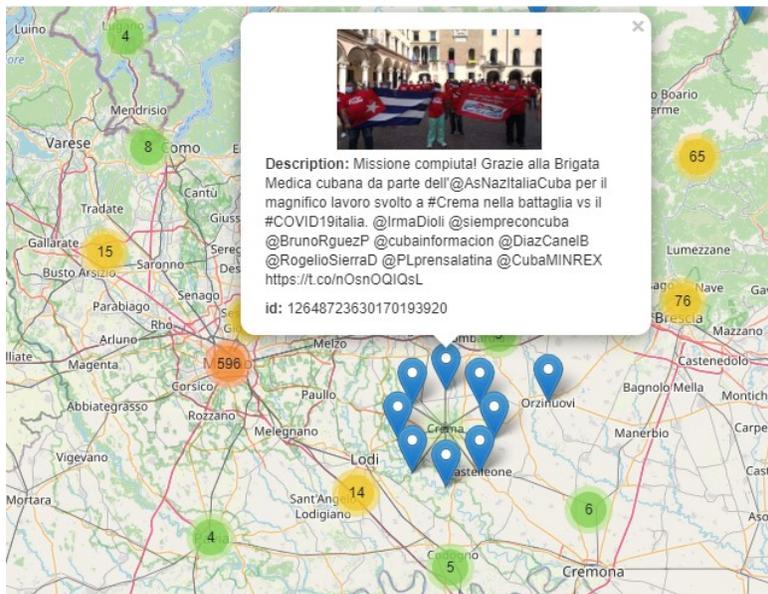


Figure 6: Visualisation of the incidents as single pin-points or clusters of adjacent points

Moreover, the platform provides the end-users with zooming, searching

and filtering functionalities in order to strengthen the exploring capabilities over the vast amount of incoming incidents/tweets (Fig. 7). Using the selection tools (polygon or circle icons) in the toolbar, a specific region of interest can be chosen on the map and only the tweets that belong to that region will be displayed, as shown in Fig. 8. Further modifications of the shape of the selected area can be carried out by using the edit layer tool.

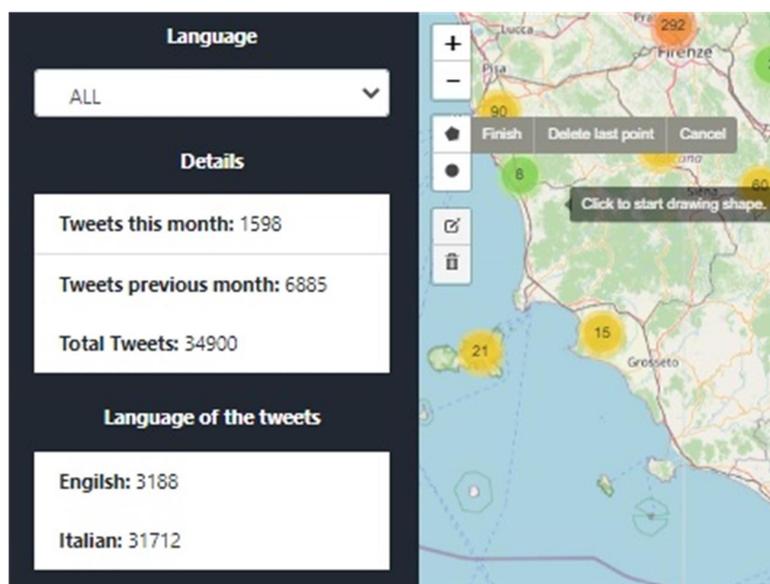


Figure 7: Details of the collection (left) and filtering, zooming and searching capabilities on the map (right)

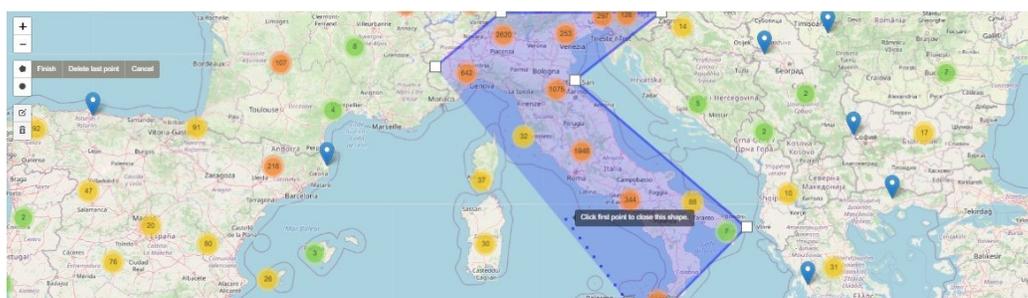


Figure 8: Filtering by selecting a region of interest on the map

The tool also enables end-users to enhance their navigation capabilities with a complementary illustration of the incidents in a table that supports

searching, sorting, paging and filtering functionalities. The columns of the table refer to *Date*, *Category*, *Latitude*, *Longitude*, *Description*, *ID*, *Reliability* and number of detected *Faces* in a post, as shown in Fig. 9.

Date	Category	Lat	Long	Description	Language	ID	Reliability	Faces
2020-07-6	covid19	34.9551	-87.2684	Ok, portiamo la mascherina. Ok, ci igienizziamo le mani. Ok, ci misuriamo la temperatura. Ok, torniamo alla normalità. Ok, non possiamo ancora togliere la mascherina. Ok, ieri sera le diapositive erano stralunate. Qui la serata con AndroidZamante. Ok. #COVID19Italia #conavirus https://t.co/F73k8Fuo08	it	1200073158284779530	0.41	186
2020-07-6	covid19	34.9551	-87.2684	Ok, portiamo la mascherina. Ok, ci igienizziamo le mani. Ok, ci misuriamo la temperatura. Ok, torniamo alla normalità. Ok, non possiamo ancora togliere la mascherina. Ok, ieri sera le diapositive erano stralunate. Qui la serata con AndroidZamante. Ok. #COVID19Italia #conavirus https://t.co/F73k8Fuo08	it	1200073294831511330	0.41	186
2020-07-6	covid19	34.9551	-87.2684	Ok, portiamo la mascherina. Ok, ci igienizziamo le mani. Ok, ci misuriamo la temperatura. Ok, torniamo alla normalità. Ok, non possiamo ancora togliere la mascherina. Ok, ieri sera le diapositive erano stralunate. Qui la serata con AndroidZamante. Ok. #COVID19Italia #conavirus https://t.co/F73k8Fuo08	it	1200073276427559840	0.41	186
2020-07-6	covid19	34.9551	-87.2684	Ok, portiamo la mascherina. Ok, ci igienizziamo le mani. Ok, ci misuriamo la temperatura. Ok, torniamo alla normalità. Ok, non possiamo ancora togliere la mascherina. Ok, ieri sera le diapositive erano stralunate. Qui la serata con AndroidZamante. Ok. #COVID19Italia #conavirus https://t.co/F73k8Fuo08	it	12000835285045248030	0.42	186
2020-07-6	covid19	34.9551	-87.2684	Ok, portiamo la mascherina. Ok, ci igienizziamo le mani. Ok, ci misuriamo la temperatura. Ok, torniamo alla normalità. Ok, non possiamo ancora togliere la mascherina. Ok, ieri sera le diapositive erano stralunate. Qui la serata con AndroidZamante. Ok. #COVID19Italia #conavirus https://t.co/F73k8Fuo08	it	1200096447029180030	0.39	186
2020-07-6	covid19	34.9551	-87.2684	Ok, portiamo la mascherina. Ok, ci igienizziamo le mani. Ok, ci misuriamo la temperatura. Ok, torniamo alla normalità. Ok, non possiamo ancora togliere la mascherina. Ok, ieri sera le diapositive erano stralunate. Qui la serata con AndroidZamante. Ok. #COVID19Italia #conavirus https://t.co/F73k8Fuo08	it	12000874675357746050	0.41	186
2020-07-6	covid19	34.9551	-87.2684	Ok, portiamo la mascherina. Ok, ci igienizziamo le mani. Ok, ci misuriamo la temperatura. Ok, torniamo alla normalità. Ok, non possiamo ancora togliere la mascherina. Ok, ieri sera le diapositive erano stralunate. Qui la serata con AndroidZamante. Ok. #COVID19Italia #conavirus https://t.co/F73k8Fuo08	it	12000874991308458250	0.41	186
2020-07-6	covid19	34.9551	-87.2684	Ok, portiamo la mascherina. Ok, ci igienizziamo le mani. Ok, ci misuriamo la temperatura. Ok, torniamo alla normalità. Ok, non possiamo ancora togliere la mascherina. Ok, ieri sera le diapositive erano stralunate. Qui la serata con AndroidZamante. Ok. #COVID19Italia #conavirus https://t.co/F73k8Fuo08	it	12000802435688872600	0.40	186
2020-07-6	covid19	34.9551	-87.2684	Ok, portiamo la mascherina. Ok, ci igienizziamo le mani. Ok, ci misuriamo la temperatura. Ok, torniamo alla normalità. Ok, non possiamo ancora togliere la mascherina. Ok, ieri sera le diapositive erano stralunate. Qui la serata con AndroidZamante. Ok. #COVID19Italia #conavirus https://t.co/F73k8Fuo08	it	1200093051122620100	0.41	186
2020-07-6	covid19	34.9551	-87.2684	Ok, portiamo la mascherina. Ok, ci igienizziamo le mani. Ok, ci misuriamo la temperatura. Ok, torniamo alla normalità. Ok, non possiamo ancora togliere la mascherina. Ok, ieri sera le diapositive erano stralunate. Qui la serata con AndroidZamante. Ok. #COVID19Italia #conavirus https://t.co/F73k8Fuo08	it	12000845700034938890	0.45	186

Figure 9: Visualisation of the incidents in a table

Regarding technical implementation, the technologies that have been applied for the development and deployment of the front-end of the Social Media Visualisation Tool include HTML, CSS, Javascript, Bootstrap, and VueJS. More specifically, the open-source JavaScript library *Leaflet*¹⁶ has been utilised to create the interactive and also mobile-friendly geographical map, the *Leaflet.markercluster*¹⁷ package for the creation of the animated and marked clusters, and *DataTables*¹⁸ for the incident table. For the back-end, nodeJS and MongoDB technologies have been employed.

4.2. Visual Analytics Dashboard

In the top right corner of the Interactive Map there is a button that navigates to the **Visual Analytics Dashboard**. This dashboard presents in a visual way the results of topic, community and event detection analysis (described in the above Sections 3.5, 3.6, and 3.7 respectively). From a menu in the left side of the page the users are able to select from a dropdown list the analytics that they wish to view. Then they can select the language of

¹⁶<https://leafletjs.com/>

¹⁷<https://github.com/Leaflet/Leaflet.markercluster>

¹⁸<https://datatables.net/>

the tweets for which the analysis will be performed, i.e. English or Italian. In community and event detection analytics there is also the option to select the date range of the publication of the tweets, while trending topics always refer to the most recent tweets and the choice of date is omitted. By clicking on the “Get” button the selected analysis runs and the results are displayed in the main component right to the menu. The visualisation of results for each analysis is described below in detail.

In topic detection (Fig. 10) each one of the detected topics is displayed as a word cloud, i.e. an illustration of single words where their importance is shown with font size. Clicking on a topic shows the complete set of tweets it comprises in a table below the word clouds, for a more low-level investigation. It should be noted that synonym words for the virus are omitted from the word clouds to improve the quality of top keywords.

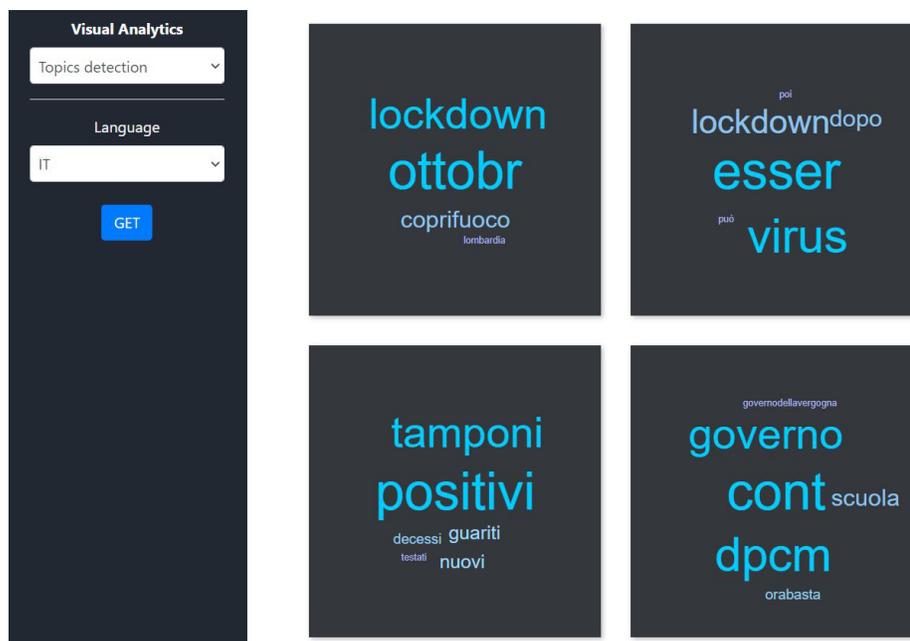


Figure 10: Detected topics as word clouds

In community detection (Fig. 11) the results are illustrated as a network graph. The graph represents the social network of Twitter accounts that are connected by mentioning one another. The unique users found in the linked pairs are displayed as nodes, pairing is displayed as edges between the two users/nodes, and communities are expressed as different colours of the nodes

and edges. The graph can also be zoomed in and out, to either take a closer look to the detected communities or view the whole picture of the network. Furthermore, a sorted list of the top ten key-players is displayed on the right of the graph. For each key-player, the community they belong to is shown, together with the number of users it comprises and coloured with the same colour of the related nodes. When clicking on the coloured text, the graph zooms in the community. Finally, when a node in the graph is clicked, then the tweets that have been posted by the respective user are displayed in a table below the graph.

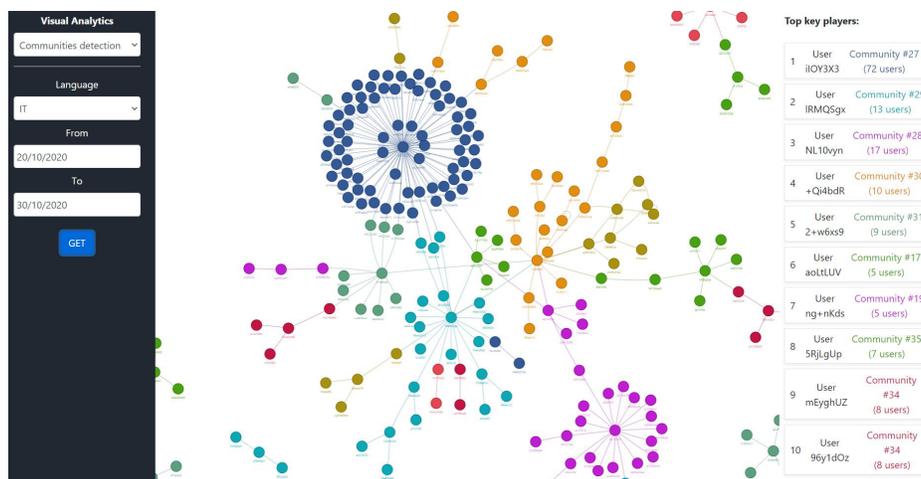


Figure 11: Visualisation of detected communities and key-players

The event detection analysis (Fig. 12) illustrates a line chart where the values in x axis represent the dates and the values in y axis the number of collected tweets per date (in certain days no tweets were collected due to technical malfunctions). Detected events are highlighted on the chart and annotated with a list of the most mentioned words (synonym words for the virus are again removed), so as to gain an insight on what that date's incident really is about.

Regarding technical implementation, some additional libraries have been used to illustrate the results of the Visual Analytics. Specifically, *Kendo*

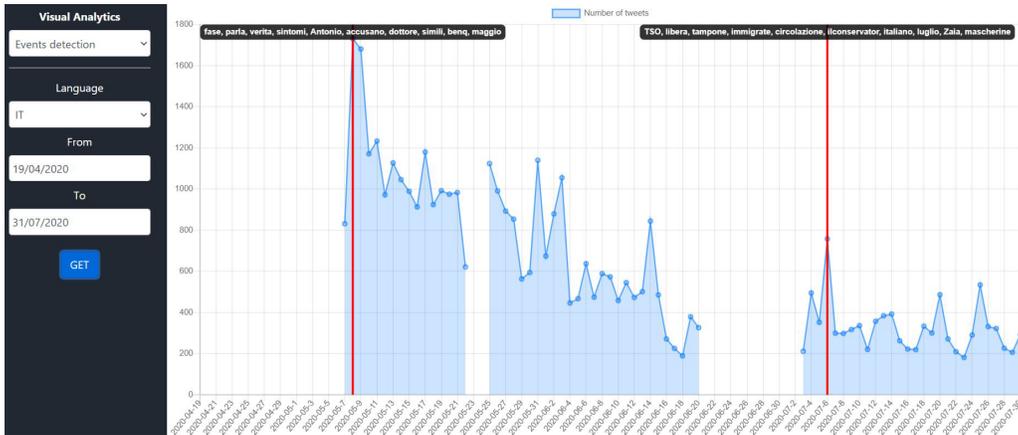


Figure 12: Number of collected tweets per day and detected events

UI^{19} and $JQCloud^{20}$ for topic detection, $vis-network^{21}$ for community detection, and $chart.js^{22}$ and its plugin $chartjs-plugin-annotation^{23}$ for the event detection.

4.3. Interaction modes

End users are able to utilise the presented online platform in order to monitor the COVID-19 situation in Italy, as expressed on social media, for alternative scopes. Specifically, visualising the automatically geotagged tweets on the Interactive Map indicates in an apparent way which areas present high activity on Twitter, thus assisting in the discovery of places at potential risk. Hence, a Health Civil Protection operator can investigate the spatial distribution of the Italian tweets in the Province of Lombardy, including the most populated cities of it. To achieve this, the utilisation of the provided tools for zooming in the map and selecting the area of interest offer a more detailed visualisation of the specific region (Fig. 13). By clicking on the pin-points the operator has access to low-level information, i.e. the original textual and visual content of single tweets.

¹⁹<https://www.telerik.com/kendo-jquery-ui>

²⁰<https://github.com/lucaong/jQCloud>

²¹<https://github.com/visjs/vis-network>

²²<https://www.chartjs.org/>

²³<https://github.com/chartjs/chartjs-plugin-annotation>

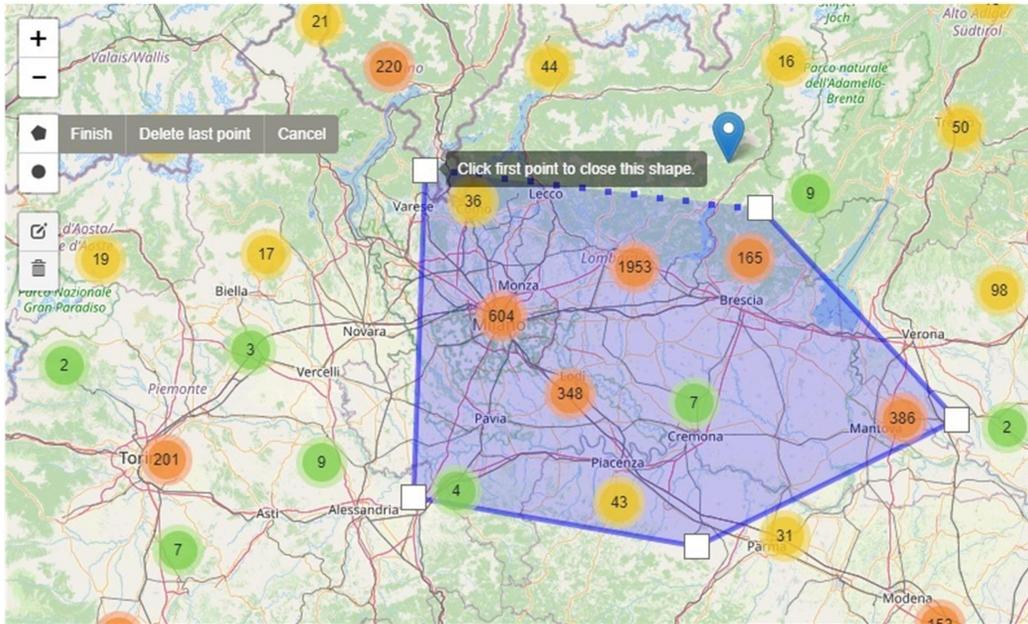


Figure 13: Distribution of Italian tweets in the Region of Lombardy

Furthermore, face detection serves as a base for crowd analysis and consequently aids in the discovery of overcrowded places by locating the human faces in the images (Fig. 4). The extracted information is useful to the operator that wants to estimate the level of compliance with the social distancing measures.

To gain deeper, high-level knowledge and better situational awareness, the health care operator can employ the provided Visual Analytics tools. Particularly, the tracking topics that are being discussed online significantly support end-users to have a better perception of the current situation, as it is affected by the virus in different ways. For example, the trending topics, as detected in the Italian Coronavirus-related Twitter posts on October 30, 2020, can be seen in Fig. 10 and involve: (a) the curfew in Lombardy, (b) the aftermath of the lockdown, (c) the new positive cases and deaths, and (d) the disapproval of the government.

The detection of online communities (Fig. 11) aids the discovery of groups of interlinked Twitter accounts that circulate fake news and conspiracies about the virus and also the identification of users (key-players) that are responsible for this harmful misinformation.

Date	Most Mentioned Keywords
08/05/2020	fase (phase), parla (speaks), verita (truth), sintomi (symptoms), Antonio, accusano (experience), dottore (doctor), simili (similar), benq (Twitter account), maggio (May)
06/07/2020	TSO (Trattamento Sanitario Obbligatorio - Mandatory Medical Treatment), libera (free), tampone (backup), immigrate (immigrant), circolazione (circulation), ilconservator (Twitter account), italiano (italian), luglio (July), Zaia (Luca Zaia), mascherine (masks)

Table 3: Most mentioned keywords for detected events

Finally, alerts from event detection, based on increasing social media activity, act as an early warning for potential outbreaks, which is very important for preventing the spread of the virus. For instance, Fig. 12 shows two discovered events, one on May 8, 2020 and one on July 6, 2020. The most mentioned keywords, which facilitate the association to the real events, can also be seen in Table 3 with their translation in English. Particularly, the first incident refers to a viral video of a doctor who states that masks cause inflammation²⁴ and the second one to the proposal of the Italian health minister for mandatory sectioning for those refusing treatment for COVID-19²⁵.

5. Evaluation

The presented work goes beyond a visualisation tool for monitoring the social media activity in relation to the COVID-19 crisis, as it constitutes an AI-enhanced system that involves state-of-the-art deep learning models, which are necessary to transform the massive streams of raw Twitter posts into meaningful and explainable information, before being visualised. The evaluation of the system focuses on the most novel of the integrated deep learning models, i.e. the localisation technique, but also concerns the face

²⁴<https://twitter.com/i/web/status/1258771433569890308>

²⁵<https://www.euroweeklynnews.com/2020/07/07/italians-who-refuse-coronavirus-treatment-could-be-sectioned/>

detection and community detection techniques, since they have been selected after a comparison of established works.

Regarding localisation, focus has been placed on the assessment of different biLSTM- and transformer-based models, considering that recent literature attested to their efficacy in sequence labeling problems, such as NER. The evaluation procedure included dataset variation testing, while the model that paired state-of-the-art results and a manageable runtime, namely the BiLSTM-CNN-CRF one, has been suggested as the most fitting for this real-time service. As far as it concerns face detection, experiments in a manually annotated collection of images from the retrieved COVID-19 set have been carried out, in order to evaluate the performance and speed of modern face detection CNNs in the task of face counting. The results indicate that the selected approach, i.e. Tinyfaces [55], yields the highest accuracy in most cases. Finally, experiments for comparing different community detection methods in regard to performance and runtime showed the superiority of the Louvain algorithm [61]. More details on the experiments are provided in the next subsections.

5.1. Evaluation of localisation

For the performance evaluation of all the localisation configurations that have been presented in Section 3.3, the values of the precision, recall and F1-score measures were computed, as well as the time the model requires to handle a relatively short collection of posts, i.e. ten COVID-19-related tweets. The latter is an important factor, since analysis needs to be completed swiftly and in real time. Moreover, since not all entities have the same significance when determining tweet locations, a decisive factor that influences the most appropriate model’s selection, is the individual LOC and ORG scores, instead of the overall F1-score for all classes. In some cases, mainly concerning the Italian language, models that presented the best overall F1-score were under-performing in the specific classes, compared to models that performed well enough in these, but not in the rest of the classes; thus, Italian-focused tests were conducted with both the default dataset configuration (4-class) and a simulated one (2-class), while the relative decision was influenced by the performance of individual scores.

English use case: Results achieved in the baseline English version of the localisation model are very similar to the state-of-the-art for English NER, yielding an F1-score of ~ 91 in the CoNLL2003 dataset. To improve on those, consequent testing was performed with the addition of BERT and

ELMo embeddings to the model pipeline, with the respective scores of the best performing models being visible in Table 4.

System (CoNLL2003)	Precision (%)	Recall (%)	F1-score (%)	Runtime (secs)
BiLSTM-CNN-CRF	89.92	91.27	90.59	8
BiLSTM-CNN-CRF + ELMo	91.94	92.90	92.42	28
BiLSTM-CNN-CRF + BERT	91.27	92.03	91.65	34
BiLSTM-CRF	90.58	91.08	90.83	8
BiLSTM-CRF + ELMo	91.69	92.99	92.33	28
BiLSTM-CRF + BERT	91.45	92.39	91.91	34
BiGRU-CRF	89.77	90.69	90.22	8
BiLSTM-CNN	88.48	90.53	89.49	6
Best shared task system: Florian et al., 2003 [64]	88.99	88.54	88.76	(not reported)
Baevski, A. et al. 2019 [65]	(not reported)	(not reported)	93.5	(not reported)

Table 4: Performance evaluation of the proposed system (EN) vs. the baseline system and state-of-the-art approaches

Both ELMo and BERT embeddings added extra efficacy to the pipeline and improved previously reported results. During tests the best scores were consistently achieved with the use of the ELMo embeddings.

Italian use case: The evaluation tests that were performed to determine which model best serves the Italian language are presented in Table 5. The evaluation findings indicate that while approaches that leverage BERT [66] did not achieve great success rates in the established EVALITA2009 dataset, they performed better in the short collection of COVID-19-related tweets. This is likely due to the specific implementation of BERT, which was pre-trained on a vocabulary of tweets, instead of prose text. Additionally, the execution time that was required to exploit the extra resources was prohibitive in a real-world scenario; the BERT-based models necessitated a considerable amount of extra time (25+ seconds), in comparison to the non-BERT ones, just to load the embeddings in system memory, and thus the specific methodology was not explored any further. Consequently, while we report in Table 5 the results for both BERT and non-BERT models, for the

actual implementation we adopt the best non-BERT approach.

System (EVALITA2009)	Precision (%)	Recall (%)	F1-score (%)	Runtime (secs)
BiLSTM-CNN-CRF (2-class)	74.73	76.66	75.68	17
BiLSTM-CNN-CRF	78.03	79.48	78.75	14
BiLSTM-CNN-CRF + BERT	72.15	74.02	73.07	46
BiLSTM-CRF	80.94	76.54	78.68	15
BiLSTM-CRF + BERT	73.27	75.47	74.36	41
BiGRU-CRF	74.79	78.21	76.46	17
BiGRU-CRF + BERT	71.06	71.55	71.30	40
BiLSTM-CNN	73.23	77.41	75.26	15
BiLSTM-CNN + BERT	65.72	72.94	69.14	25
DNN: Basile, Semeraro, & Cassotti, 2017 [67]	82.86	81.82	82.34	(not reported)
Best shared task system: FBK_ZanoliPianta [68]	84.07	80.02	82.00	(not reported)
Rerank model: Nguyen and Moschitti, 2012 [69]	85.99	82.73	84.33	(not reported)

Table 5: Performance evaluation of the proposed system (IT) vs. the baseline system and state-of-the-art approaches

Additionally, tests were also performed with a modified version of the dataset, where the LOC and GPE classes were unified into one, the ORG one remained the same and the PER class was completely omitted. The 2-class accumulated results did not present significant alterations to the ones of the full dataset, with 2% improvement in the ORG class (69.4% 2-class vs 67.7% full) and 2% deterioration in the LOC joint class (79.88% 2-class vs 81.95% full). However, recognition results were significantly ameliorated when testing the model on the COVID-19 example tweets, with previously misclassified entities now receiving the correct annotation.

According to the results, the application of additional linguistic resources, such as contextual embeddings, favours the selected model, rendering it capable of achieving better recognition results. However, there is a toll in computation efficiency, since the reported runtimes are increased respectively. Hence, to manage an almost real-time processing of COVID-19-related tweets, localisation needs to be based on a fast biLSTM implementation. The

BiLSTM-CNN-CRF approach was the most appropriate candidate for the selected task in both languages, combining great F1 results with manageable processing time.

5.2. Evaluation of face detection

In this section, the evaluation of a selection of state-of-the-art face detection methods is presented, as a means of providing motivation for the choice of integrating the Tinyfaces detector [55] to the proposed framework.

In order to compile the evaluation dataset, a set of 100 Covid19-related images were manually selected and annotated, with the total count of faces a human eye can observe in each image. The manual annotation process was carried out by two individuals who were given the instruction of adding a face to the total count, for every patch in the image that can be recognized as a human face or a facial part. The evaluation dataset contains images with several key challenges, such as crowded scenes with small scale faces and mask-wearing persons with partially occluded faces. Note that human annotators were only used so as to construct the test set upon which the evaluation of automatic face detectors was performed.

Huang et al. presented in [70] an analysis on the accuracy and speed trade-offs of modern convolutional generic object detectors. In order to provide a small benchmark, we select the most accurate and the fastest from the pool of recent state-of-the-art meta-architecture and feature extractor combinations in [70], and test them for the task of counting human faces in Covid19-related images. The fastest one, uses the SSD architecture which is a single feed-forward convolutional network that predicts object classes and anchor offsets. It has been paired with the updated MobileNetV2 backbone feature extractor, presented in [71]. The most accurate in [70] was found to be the Inception Resnet V2 feature extractor coupled with the Faster R-CNN architecture, which deploys a region proposal network, to detect class-agnostic object proposals first, and then a box classifier network, so as to assign classes and refine box coordinates. Both models have been pre-trained using the Open Images V4 dataset [72], and can be acquired as off-the-shelf models from the Tensorflow repository [73].

The metrics that are used to evaluate the performance of the models on face counting are the Mean Squared Error (MSE) and Mean Absolute Error (MAE):

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2,$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i|,$$

where, Y_i is the true (annotated) number of faces and \hat{Y}_i is the number of detected faces, in image i , when i is ranging from 1 to n . Moreover, the running time is given for each model in frames per second, so as to assess the efficiency of each one.

Model	MSE	MAE	fps
Faster-RCNN with Inception Resnet V2	26.91	1.45	1.53
SSD with MobileNetV2	31.53	1.99	6.63
Tinyfaces	5.35	0.61	0.47

Table 6: Evaluation of face detectors for the task of face counting.

Table 6 shows the evaluation results, which were calculated after passing the annotated images through the 3 selected face detectors. It can be seen that, Tinyfaces performs better than the Faster R-CNN and the SSD architectures judging by both the average squared and absolute errors. Specifically, Tinyfaces greatly outperforms the other two in the MSE metric, which greatly penalizes large errors. This may indicate that overcrowded scenes with a high number of faces cannot be detected well by the other two models. In terms of speed, Tinyfaces has the slowest response, but given the high accuracy of predictions in the critical crowded scenes, the overhead is acceptable.

5.3. Evaluation of community detection

This section presents the experiment realised towards comparing different community detection approaches in terms of performance (modularity) and execution time, in order to find the most appropriate algorithm and adopt it in our methodology (Section 3.6). The examined approaches are the Edge Betweenness [74], Fast Greedy [75], Label Propagation [76], Louvain [61], Walktrap [77], and Infomap [78] [79] algorithms, all implemented in R language. The dataset selected for this experiment comprises tweets about COVID-19 in Italy, which have been collected between 03/07/2020 and 17/07/2020; a time period in which there was plenty of discussion online about a proposal for mandatory medical treatment.

The first metric to be investigated is modularity, a measure of the structure of networks. It has been designed to measure the strength of division of a network into modules, i.e. communities. Networks with high modularity have dense connections between the nodes within modules, but sparse connections between nodes in different modules. We aim at the maximisation of modularity, defined as:

$$Q = \frac{1}{2m} \sum_{i=1}^c (e_{ii} - a_i^2)$$

where e_{ij} is the fraction of links between a node in community i and a node in community j , a_i is the fraction of links between two members of the community i , $m = \sum_k deg(n_k)$ and c is the number of communities. Fig. 14 shows the results of modularity and Louvain appears to outperform the other methods.

The second metric that is examined is the execution time of each algorithm and the results are depicted in Fig. 15. It has to be noted that Edge Betweenness is excluded from the figure, because it is extremely slower than the others (with an average execution time of 12.25 seconds) and it would affect the readability of the figure. Among the other five algorithms, Louvain and Label Propagation are proven to be the fastest and most scalable options. This, in combination with the highest achieved modularity, signifies the superiority of the Louvain algorithm compared to other methods and justifies our selection to adopt it in our methodology.

6. Conclusions

In this paper we presented an integrated system that collects Twitter posts in a real-time manner, analyses and visualises them in a multi-level aspect. Particularly, low-level knowledge is extracted by the analysis of the tweets to estimate the reliability of the information they bear, to assign coordinates based on the locations that are mentioned in their text, and to detect the number of people appearing in their visual content, and the tweets are visualised as pin-points on an interactive map-based platform. Furthermore, high-level knowledge is discovered by the aggregated analysis of the obtained tweets. Therefore, methodologies for event detection by monitoring the remarkable differences in the daily distribution of the number of tweets, community detection with identification of the most influential



Figure 14: Modularity of communities detected by each algorithm as days increase

users, and topic detection by clustering textual streams with similar content have been encapsulated in the proposed unified framework, accompanied by the visualisations of their outcomes.

The proposed framework has been applied to monitor tweets about COVID-19 in the area of Italy, by defining as search keywords the trending hashtags “#COVID19Italy” and “#COVID19Italia”, and more than 40,000 tweets have been collected in the first three months of application. The evaluation of the knowledge extraction methodologies were promising and encouraging to further continue in this direction. Regarding the automatic geotagging, a (biLSTM)-based model, specifically trained for the English and Italian languages, was evaluated on the CoNLL2003 dataset for English and on the Evalita2009 dataset for Italian. The best models were selected based on a combination of best F1-scores and runtime; for English the best model achieved a 90.59 F1-score/8 secs result, while the respective Italian one reached a score of 78.75 F1-score/14 secs. As far as it concerns the face detection approach, the Tinyfaces methodology was deployed and evaluated

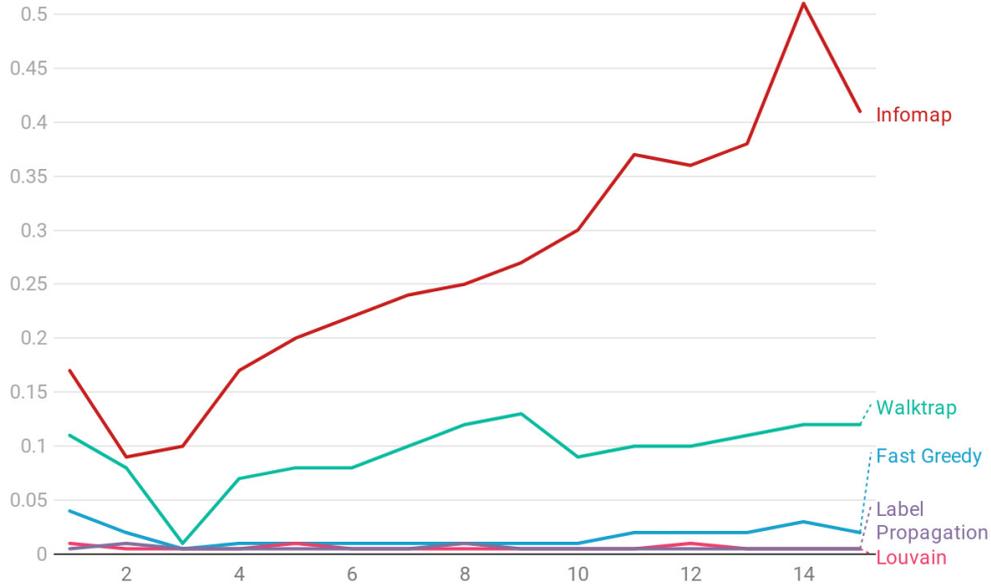


Figure 15: Execution time (in seconds) of each community detection algorithm as days increase

for the task of face counting, along with other SoA detectors. The results indicate better performance on small scale faces in highly crowded scenes. Finally, the comparison of alternative community detection techniques by means of modularity and execution time showed the superiority of the Louvain algorithm that is adopted in our solution.

Generally, this work aims to support authorities and organisations from the health sector or civil protection in the management of the pandemic crisis. The visualisation of tweets on a map indicates areas with high activity, while the filtering/search capabilities enable the discovery of certain types of posts. The estimation of the number of faces in Twitter images can identify cases of overcrowding, the detection of recent topics can provide awareness of the issues that concern the society, the analysis of user communities can expose accounts responsible for misinformation and event detection can contribute with alerts of possible outbreaks.

Future directions for the system include the maintenance of a low response time of showing the results as the collection steadily grows, the development of more advanced filters in the Interactive Map, and the extension of the

Visual Analytics Dashboard with further methodologies that assist the situational awareness of end-users and their response to the ongoing pandemic crisis. Towards this direction, a consolidation with quantitative and qualitative information, such as demographic and statistical information related to the COVID-19 pandemic crisis in the region of interest could be useful to enhance the assessment and awareness of the situation by stakeholders and assist them in the effective decision-making process.

Acknowledgements

This work was supported by the EU's Horizon 2020 research and innovation programme under grant agreements H2020-883484 PathoCERT, H2020-883284 7SHIELD and H2020-832876 aqua3S.

References

- [1] L. M. Aiello, G. Petkos, C. Martin, D. Corney, S. Papadopoulos, R. Skraba, A. Göker, I. Kompatsiaris, A. Jaimes, Sensing trending topics in twitter, *IEEE Transactions on Multimedia* 15 (6) (2013) 1268–1282.
- [2] C. S. Lee, L. Ma, News sharing in social media: The effect of gratifications and prior experience, *Computers in human behavior* 28 (2) (2012) 331–339.
- [3] Y. Kryvasheyeu, H. Chen, N. Obradovich, E. Moro, P. Van Hentenryck, J. Fowler, M. Cebrian, Rapid assessment of disaster damage using social media activity, *Science advances* 2 (3) (2016) e1500779.
- [4] H. Gao, G. Barbier, R. Goolsby, Harnessing the crowdsourcing power of social media for disaster relief, *IEEE Intelligent Systems* 26 (3) (2011) 10–14.
- [5] Z. Wang, X. Ye, Social media analytics for natural disaster management, *International Journal of Geographical Information Science* 32 (1) (2018) 49–72.
- [6] Y. Xie, Z. Chen, Y. Cheng, K. Zhang, A. Agrawal, W.-K. Liao, A. Choudhary, Detecting and tracking disease outbreaks by mining social media data, in: *Twenty-Third International Joint Conference on Artificial Intelligence*, 2013.

- [7] A. Sadilek, H. Kautz, V. Silenzio, Modeling spread of disease from social interactions, in: Sixth International AAAI Conference on Weblogs and Social Media, 2012.
- [8] L. E. Charles-Smith, T. L. Reynolds, M. A. Cameron, M. Conway, E. H. Lau, J. M. Olsen, J. A. Pavlin, M. Shigematsu, L. C. Streichert, K. J. Suda, et al., Using social media for actionable disease surveillance and outbreak management: a systematic literature review, *PloS one* 10 (10) (2015) e0139701.
- [9] M. Cinelli, W. Quattrociocchi, A. Galeazzi, C. M. Valensise, E. Brugnoli, A. L. Schmidt, P. Zola, F. Zollo, A. Scala, The covid-19 social media infodemic, *arXiv preprint arXiv:2003.05004* (2020).
- [10] T. Alshaabi, J. R. Minot, M. V. Arnold, J. L. Adams, D. R. Dewhurst, A. J. Reagan, R. Muhamad, C. M. Danforth, P. S. Dodds, How the world’s collective attention is being paid to a pandemic: Covid-19 related 1-gram time series for 24 languages on twitter, *arXiv preprint arXiv:2003.12614* (2020).
- [11] S. R. Rufai, C. Bunce, World leaders’ usage of twitter in response to the covid-19 pandemic: a content analysis, *Journal of Public Health* (2020).
- [12] Y. Zhao, H. Xu, Chinese public attention to covid-19 epidemic: Based on social media, *medRxiv* (2020).
- [13] A. Abd-Alrazaq, D. Alhuwail, M. Househ, M. Hamdi, Z. Shah, Top concerns of tweeters during the covid-19 pandemic: infoveillance study, *Journal of medical Internet research* 22 (4) (2020) e19016.
- [14] E. T. K. Sang, M. Schraagen, M. Dastani, S. Wang, Discovering pandemic topics on twitter (2021).
- [15] Y. Su, A. Venkat, Y. Yadav, L. B. Puglisi, S. J. Fodeh, Twitter-based analysis reveals differential covid-19 concerns across areas with socioeconomic disparities, *Computers in Biology and Medicine* (2021) 104336.
- [16] H. W. Park, S. Park, M. Chong, Conversations and medical news frames on twitter: Infodemiological study on covid-19 in south korea, *Journal of Medical Internet Research* 22 (5) (2020) e18897.

- [17] Y. Bang, E. Ishii, S. Cahyawijaya, Z. Ji, P. Fung, Model generalization on covid-19 fake news detection, arXiv preprint arXiv:2101.03841 (2021).
- [18] W. S. Paka, R. Bansal, A. Kaushik, S. Sengupta, T. Chakraborty, Cross-sean: A cross-stitch semi-supervised neural attention model for covid-19 fake news detection, arXiv preprint arXiv:2102.08924 (2021).
- [19] K. Sharma, S. Seo, C. Meng, S. Rambhatla, Y. Liu, Covid-19 on social media: Analyzing misinformation in twitter conversations, arXiv preprint arXiv:2003.12309 (2020).
- [20] R. Kouzy, J. Abi Jaoude, A. Kraitem, M. B. El Alam, B. Karam, E. Adib, J. Zarka, C. Traboulsi, E. W. Akl, K. Baddour, Coronavirus goes viral: quantifying the covid-19 misinformation epidemic on twitter, *Cureus* 12 (3) (2020).
- [21] J. S. Brennen, F. M. Simon, P. N. Howard, R. K. Nielsen, Types, sources, and claims of covid-19 misinformation, Reuters Institute (2020).
- [22] E. Ferrara, What types of covid-19 conspiracies are populated by twitter bots?, *First Monday* (2020).
- [23] K.-C. Yang, C. Torres-Lugo, F. Menczer, Prevalence of low-credibility information on twitter during the covid-19 outbreak, arXiv preprint arXiv:2004.14484 (2020).
- [24] L. Singh, S. Bansal, L. Bode, C. Budak, G. Chi, K. Kawintiranon, C. Padden, R. Vanarsdall, E. Vraga, Y. Wang, A first look at covid-19 information and misinformation sharing on twitter, arXiv preprint arXiv:2003.13907 (2020).
- [25] W. Ahmed, J. Vidal-Alaball, J. Downing, F. L. Seguí, Covid-19 and the 5g conspiracy theory: social network analysis of twitter data, *Journal of Medical Internet Research* 22 (5) (2020) e19458.
- [26] C. Li, L. J. Chen, X. Chen, M. Zhang, C. P. Pang, H. Chen, Retrospective analysis of the possibility of predicting the covid-19 outbreak from internet searches and social media data, china, 2020, *Eurosurveillance* 25 (10) (2020) 2000199.

- [27] K. Jahanbin, V. Rahmanian, Using twitter and web news mining to predict covid-19 outbreak, *Asian Pacific Journal of Tropical Medicine* 13 (2020).
- [28] E. Chen, K. Lerman, E. Ferrara, Tracking social media discourse about the covid-19 pandemic: Development of a public coronavirus twitter data set, *JMIR Public Health and Surveillance* 6 (2) (2020) e19273.
- [29] S. Alqurashi, A. Alhindi, E. Alanazi, Large arabic twitter dataset on covid-19, *arXiv preprint arXiv:2004.04315* (2020).
- [30] C. E. Lopez, M. Vasu, C. Gallemore, Understanding the perception of covid-19 policies by mining a multilanguage twitter dataset, *arXiv preprint arXiv:2003.10359* (2020).
- [31] J. M. Banda, R. Tekumalla, G. Wang, J. Yu, T. Liu, Y. Ding, G. Chowell, A large-scale covid-19 twitter chatter dataset for open scientific research—an international collaboration, *arXiv preprint arXiv:2004.03688* (2020).
- [32] M. Y. Kabir, S. Madria, Coronavis: A real-time covid-19 tweets data analyzer and data repository (2020).
- [33] U. Naseem, I. Razzak, M. Khushi, P. W. Eklund, J. Kim, Covidsentiment: A large-scale benchmark twitter data set for covid-19 sentiment analysis, *IEEE Transactions on Computational Social Systems* (2021).
- [34] D. Dimitrov, E. Baran, P. Fafalios, R. Yu, X. Zhu, M. Zloch, S. Dietze, Tweetscov19—a knowledge base of semantically annotated tweets about the covid-19 pandemic, *arXiv preprint arXiv:2006.14492* (2020).
- [35] T. Yang, K. Shen, S. He, E. Li, P. Sun, P. Chen, L. Zuo, J. Hu, Y. Mo, W. Zhang, H. Zhang, J. Chen, Y. Guo, Covidnet: To bring data transparency in the era of covid-19 (2020). *arXiv:2005.10948*.
- [36] OmniSci, Omnisci technical whitepaper - big data analytics at speed and scale, OmniSci Inc. (2018).
- [37] C. Boididou, S. Papadopoulos, L. Apostolidis, Y. Kompatsiaris, Learning to detect misleading content on twitter, in: *Proceedings of the 2017 ACM on International Conference on Multimedia Retrieval, 2017*, pp. 278–286.

- [38] C. Boididou, S. Papadopoulos, M. Zampoglou, L. Apostolidis, O. Papadopoulou, Y. Kompatsiaris, Detection and visualization of misleading content on twitter, *International Journal of Multimedia Information Retrieval* 7 (1) (2018) 71–86.
- [39] G. Lample, M. Ballesteros, S. Subramanian, K. Kawakami, C. Dyer, Neural architectures for named entity recognition, *arXiv preprint arXiv:1603.01360* (2016).
- [40] M. E. Peters, W. Ammar, C. Bhagavatula, R. Power, Semi-supervised sequence tagging with bidirectional language models, *arXiv preprint arXiv:1705.00108* (2017).
- [41] X. Ma, E. Hovy, End-to-end sequence labeling via bi-directional lstm-cnns-crf, *arXiv preprint arXiv:1603.01354* (2016).
- [42] J. P. Chiu, E. Nichols, Named entity recognition with bidirectional lstm-cnns, *Transactions of the Association for Computational Linguistics* 4 (2016) 357–370.
- [43] E. F. Sang, F. De Meulder, Introduction to the conll-2003 shared task: Language-independent named entity recognition, *arXiv preprint cs/0306050* (2003).
- [44] M. Speranza, The named entity recognition task at evalita 2009, in: *EVALITA 2009*, 2009.
- [45] N. Reimers, I. Gurevych, Optimal hyperparameters for deep lstm-networks for sequence labeling tasks, *arXiv preprint arXiv:1707.06799* (2017).
- [46] J. Pennington, R. Socher, C. D. Manning, Glove: Global vectors for word representation, in: *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, 2014, pp. 1532–1543.
- [47] E. Grave, P. Bojanowski, P. Gupta, A. Joulin, T. Mikolov, Learning word vectors for 157 languages, *arXiv preprint arXiv:1802.06893* (2018).
- [48] P. Viola, M. Jones, Rapid object detection using a boosted cascade of simple features, in: *Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001*, Vol. 1, IEEE, 2001, pp. I–I.

- [49] C. Shu, X. Ding, C. Fang, Histogram of the oriented gradient for face recognition, *Tsinghua Science and Technology* 16 (2) (2011) 216–224.
- [50] T. Ahonen, J. Matas, C. He, M. Pietikäinen, Rotation invariant image description with local binary pattern histogram fourier features, in: *Scandinavian conference on image analysis*, Springer, 2009, pp. 61–70.
- [51] S. Zhang, X. Zhu, Z. Lei, H. Shi, X. Wang, S. Z. Li, S3fd: Single shot scale-invariant face detector, in: *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 192–201.
- [52] X. Sun, P. Wu, S. C. Hoi, Face detection using deep learning: An improved faster rcnn approach, *Neurocomputing* 299 (2018) 42–50.
- [53] J. Li, Y. Wang, C. Wang, Y. Tai, J. Qian, J. Yang, C. Wang, J. Li, F. Huang, Dsfd: dual shot face detector, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 5060–5069.
- [54] J. Deng, J. Guo, E. Ververas, I. Kotsia, S. Zafeiriou, Retinaface: Single-shot multi-level face localisation in the wild, in: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 5203–5212.
- [55] P. Hu, D. Ramanan, Finding tiny faces, in: *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 951–959.
- [56] S. Yang, P. Luo, C.-C. Loy, X. Tang, Wider face: A face detection benchmark, in: *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 5525–5533.
- [57] I. Gialampoukidis, S. Vrochidis, I. Kompatsiaris, A hybrid framework for news clustering based on the dbscan-martingale and lda, in: *International Conference on Machine Learning and Data Mining in Pattern Recognition*, Springer, 2016, pp. 170–184.
- [58] D. M. Blei, A. Y. Ng, M. I. Jordan, Latent dirichlet allocation, *Journal of machine Learning research* 3 (Jan) (2003) 993–1022.

- [59] I. Gialampoukidis, S. Vrochidis, I. Kompatsiaris, I. Antoniou, Probabilistic density-based estimation of the number of clusters using the dbscan-martingale process, *Pattern Recognition Letters* 123 (2019) 23–30.
- [60] I. Gialampoukidis, G. Kalpakis, T. Tsirikika, S. Papadopoulos, S. Vrochidis, I. Kompatsiaris, Detection of terrorism-related twitter communities using centrality scores, in: *Proceedings of the 2nd International Workshop on Multimedia Forensics and Security*, 2017, pp. 21–25.
- [61] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, E. Lefebvre, Fast unfolding of communities in large networks, *Journal of statistical mechanics: theory and experiment* 2008 (10) (2008) P10008.
- [62] I. Gialampoukidis, G. Kalpakis, T. Tsirikika, S. Vrochidis, I. Kompatsiaris, Key player identification in terrorism-related social media networks using centrality measures, in: *2016 European Intelligence and Security Informatics Conference (EISIC)*, IEEE, 2016, pp. 112–115.
- [63] R. E. Shiffler, Maximum z scores and outliers, *The American Statistician* 42 (1) (1988) 79–80.
- [64] R. Florian, A. Ittycheriah, H. Jing, T. Zhang, Named entity recognition through classifier combination, in: *Proceedings of the seventh conference on Natural language learning at HLT-NAACL 2003*, 2003, pp. 168–171.
- [65] A. Baevski, S. Edunov, Y. Liu, L. Zettlemoyer, M. Auli, Cloze-driven pretraining of self-attention networks, *arXiv preprint arXiv:1903.07785* (2019).
- [66] M. Polignano, P. Basile, M. de Gemmis, G. Semeraro, V. Basile, AIBERTO: Italian BERT Language Understanding Model for NLP Challenging Tasks Based on Tweets, in: *Proceedings of the Sixth Italian Conference on Computational Linguistics (CLiC-it 2019)*, Vol. 2481, CEUR, 2019.
 URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85074851349&partnerID=40&md5=7abed946e06f76b3825ae5e294ffac14>

- [67] P. Basile, G. Semeraro, P. Cassotti, Bi-directional lstm-cnns-crf for italian sequence labeling, CLiC-it 2017 11-12 December 2017, Rome (2017) 18.
- [68] R. Zanolì, E. Pianta, C. Giuliano, Named entity recognition through redundancy driven classifiers, Proceedings of EVALITA 9 (2009).
- [69] T.-V. T. Nguyen, A. Moschitti, Structural reranking models for named entity recognition, *Intelligenza Artificiale* 6 (2) (2012) 177–190.
- [70] J. Huang, V. Rathod, C. Sun, M. Zhu, A. Korattikara, A. Fathi, I. Fischer, Z. Wojna, Y. Song, S. Guadarrama, et al., Speed/accuracy trade-offs for modern convolutional object detectors, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 7310–7311.
- [71] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, L.-C. Chen, Mobilenetv2: Inverted residuals and linear bottlenecks, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 4510–4520.
- [72] A. Kuznetsova, H. Rom, N. Alldrin, J. Uijlings, I. Krasin, J. Pont-Tuset, S. Kamali, S. Popov, M. Mallocci, T. Duerig, et al., The open images dataset v4: Unified image classification, object detection, and visual relationship detection at scale, arXiv preprint arXiv:1811.00982 (2018).
- [73] J. Huang, V. Rathod, Supercharge your computer vision models with the tensorflow object detection api (2017).
URL <https://ai.googleblog.com/2017/06/supercharge-your-computer-vision-models.html>
- [74] M. E. Newman, M. Girvan, Finding and evaluating community structure in networks, *Physical review E* 69 (2) (2004) 026113.
- [75] A. Clauset, M. E. Newman, C. Moore, Finding community structure in very large networks, *Physical review E* 70 (6) (2004) 066111.
- [76] U. N. Raghavan, R. Albert, S. Kumara, Near linear time algorithm to detect community structures in large-scale networks, *Physical review E* 76 (3) (2007) 036106.

- [77] P. Pons, M. Latapy, Computing communities in large networks using random walks, in: International symposium on computer and information sciences, Springer, 2005, pp. 284–293.
- [78] M. Rosvall, D. Axelsson, C. T. Bergstrom, The map equation, The European Physical Journal Special Topics 178 (1) (2009) 13–23.
- [79] L. Bohlin, D. Edler, A. Lancichinetti, M. Rosvall, Community detection and visualization of networks with the map equation framework, in: Measuring scholarly impact, Springer, 2014, pp. 3–34.