Assessing vulnerability to psychological distress during the COVID-19 pandemic through the analysis of microblogging content

Marco Viviani [a,*], Cristina Crocamo [b], Matteo Mazzola [a], Francesco Bartoli [b,c], Giuseppe Carrà [b,c,d,1], Gabriella Pasi [a,1]

[a] *Department of Informatics, Systems, and Communication (DISCo), University of Milano-Bicocca, Milan, Italy*
[b] *Department of Medicine and Surgery, University of Milano-Bicocca, Monza, Italy*
[c] *Department of Mental Health & Addiction, ASST Nord Milano, Bassini Hospital, Cinisello Balsamo, Italy*
[d] *Division of Psychiatry, University College London (UCL), London, UK*

ARTICLE INFO

ABSTRACT

In recent years we have witnessed a growing interest in the analysis of social media data under different perspectives, since these online platforms have become the preferred tool for generating and sharing content across different users organized into virtual communities, based on their common interests, needs, and perceptions. In the current study, by considering a collection of social textual contents related to COVID-19 gathered on the Twitter microblogging platform in the period between August and December 2020, we aimed at evaluating the possible effects of some critical factors related to the pandemic on the mental well-being of the population. In particular, we aimed at investigating potential lexicon identifiers of vulnerability to psychological distress in digital social interactions with respect to distinct COVID-related scenarios, which could be "at risk" from a psychological discomfort point of view. Such scenarios have been associated with peculiar topics discussed on Twitter. For this purpose, two approaches based on a "top-down" and a "bottom-up" strategy were adopted. In the top-down approach, three potential scenarios were initially selected by medical experts, and associated with topics extracted from the Twitter dataset in a hybrid unsupervised-supervised way. On the other hand, in the bottom-up approach, three topics were extracted in a totally unsupervised way capitalizing on a Twitter dataset filtered according to the presence of keywords related to vulnerability to psychological distress, and associated with at-risk scenarios. The identification of such scenarios with both approaches made it possible to capture and analyze the potential psychological vulnerability in critical situations.

© 2021 Published by Elsevier B.V.

1. Introduction

Today, social media have become the preferred tool for generating and sharing content with friends or unknown people, based on common interests, needs, and perceptions. In particular, microblogging sites are especially geared toward the exchange of textual content and enable to build virtual communities around specific *topics* of interest, giving rise to *conversations* that often refer to real-life events and scenarios.

Considering such relevant topics, it is possible to build the so-called *conversation graphs*, i.e., specific network-based structures where nodes represent individuals discussing a specific topic, and edges show different types of social interactions among them. These structures can be analyzed at different levels, including a *topology-based* one, i.e., to study the peculiar interactions and the establishment of communities among users, and a *content-based* perspective, i.e., to study the lexical, semantic and sentiment-related aspects inherent the shared texts within the community.

In social computing and related research fields, several approaches have analyzed microblogging platforms and conversation graphs built on top of them, ranging from the study of *opinion polarization* and the consequent formation of *echo chambers* in the political debate [1,2], to *rumor detection* in the spread of health-related information [3,4], just to mention the most recent application scenarios. In other recent works related to the healthcare and public health domain, the value of considering microblogging platforms for a better understanding of mental health states is particularly marked, given that it provides access to individual accounts of user behaviors, activities, thoughts, and feelings that may be indicative of emotional well-being [5].

---

* Corresponding author.

E-mail addresses: marco.viviani@unimib.it (M. Viviani), cristina.crocamo@unimib.it (C. Crocamo), m.mazzola29@campus.unimib.it (M. Mazzola), francesco.bartoli@unimib.it (F. Bartoli), giuseppe.carra@unimib.it (G. Carrà), gabriella.pasi@unimib.it (G. Pasi).
[1] Equal contribution.

In the current study, we aim to consider the social content shared on Twitter and the interactions among individuals related to the COVID-19 pandemic, pursuing the identification of potential "at-risk" scenarios from the psychological distress point of view. In particular, we use a two-fold approach. In a first approach, defined as "top-down", three potential scenarios are provided according to the specific expertise of the medical team, which are associated with specific topics in Twitter by means of a hybrid unsupervised-supervised technique. Subsequently, the conversation graphs resulting from interactions between users' tweets related to those topics are analyzed from the point of view of the presence of potential lexicon identifiers related to vulnerability to psychological distress. In a second approach, defined as "bottom-up", tweets encompassing terms related to negative emotions and distress are first identified and then, through a totally unsupervised approach, emergent topics, possibly mirroring at-risk scenarios, are automatically generated and analyzed.

From the outcomes of the analyses provided in the article with respect to both approaches, interesting results have been obtained that allow us to qualitatively assess (thanks to the team of medical experts) the degree of vulnerability to psychological distress relating to the considered scenarios, being also able to capture distinct categories of psychological and emotional states related to the language employed.

2. Background and related work

In recent years, there has been a growing interest in the study of social media data for public health concerns, enabling to study the domain of vulnerability to psychological distress by exploring features that are hardly detectable by means of classical epidemiological designs [6,7]. Research focused largely on the analysis of *User-Generated Content* (UGC) that shape social media data, encompassing the psychological well-being of individuals and populations according to early warning indicators of emotional resilience [8].

Indeed, the language and patterns of communication on social media seem providing multifaceted community-level information possibly related to plausible indicators of psychosocial health that may complement other traditional methods [9]. Prior studies leveraged social media and explored their utility to better understand, identify, and characterize mental health-related conditions, monitoring attitudes, including problematic Web use [10–12], and examining proxy variables of psychopathological conditions potentially associated with health outcomes [13–16]. In particular, De Choudhury and colleagues [13], using crowdsourcing data generation methods and machine learning jointly, found that decreased social activity, increased negative affect, highly-clustered ego networks, and increased relational and medical concerns, were all related to depressive symptoms. A more recent work has been proposed with the aim of facilitating the identification of depressive symptoms from Twitter data, based on natural language processing techniques. In particular, an annotated corpus of Twitter posts was developed, also thanks to external depression-related lexicons, to enable a better understanding of conveyed depressive symptoms and psychosocial stressors [14]. Additional research focused on concerns, opinions, and perceptions embedded in Twitter posts, including emotional resilience [17–20]. Originating sentiment trajectories on social media might represent a novel potential approach to monitor and support emotional wellbeing at a community-level [21–23].

Many of the above and other works in the literature, have considered the Twitter microblogging platform and the analysis of tweets content to develop powerful methods (both supervised through the use of experts or external resources, and unsupervised or hybrid) to explore sentiment, shared topics, themes,

and sources. Different techniques have been exploited to identify potential clusters according to individuals' personal characteristics and circumstances [24]. By using both classification (i.e., supervised) and clustering (i.e., unsupervised) algorithms, changes in user opinions and perceptions over time and across different regions were tracked and a moderate strength of relationship between exposure to social media content and individual perceptions was detected [25].

Previous evidence showed how digital platforms and social media like Twitter can be leveraged for behavioral modeling and representation, providing an opportunity to better understand health-related concerns and the mechanisms of social influence that may drive different behavioral responses, and monitor inter- and intra-personal psychosocial mediators [26]. Thus, Twitter appeared as a promising surveillance tool for public health preparedness, response, and recovery, enabling to combine context analysis and the semantic spectrum of User-Generated Content for the identification of type of communication and related sentiment and emotions.

This was true also considering the recent attempt to control the spread of epidemic diseases, such as COVID-19, and the widespread physical distancing limiting informal social interactions [27]. For instance, a study of 580 million tweets posted during the early months (i.e., January to May 2020) of the COVID-19 pandemic used the geographic information associated with tweets as a proxy for human mobility assessing adherence to guidelines about physical distancing [28]. In addition, a social media debate, with subjective perspectives mixed-up with legitimate and authoritative sources of information, emerged during the fight against the COVID-19 pandemic possibly resulting in emotional contagion.

The online spread of emotion-related content, leading people to experience the same emotions often without their awareness, suggested the need for novel surveillance approaches based on timely evaluations of sentiment trajectories [23]. Such trajectories were assessed using both lexicon-based methods and properly trained and fine-tuned semantic-based models for COVID-19 related tweets, actually capturing variations also in terms of emotional contagion [20,22,23,29,30]. Moreover, stress, anxiety, and loneliness levels detected in COVID-19 related tweets in 2020 seemed increasingly divergent from 2019 ones, providing insight into likely changes in mental health of communities and early recognition of hot-spots of declining mental health [31]. Relevant findings advocated the potential that social media like Twitter offer about the monitoring of the emotional and psychological aspects of individual perceptions during the pandemic waves, entailing appropriate preventive activities via social media as part of the public health response to COVID-19.

Hence, social media platforms might constitute a promising environment for identifying potential digital identifiers of vulnerability to psychological distress [32]. This may suggest that, by taking into account specific at-risk scenarios, which we correspond to conversational graphs built around (emergent) discussion topics, virtual communities can be followed-up with respect to topic-dependent psychological vulnerability.

3. A methodological solution for analyzing vulnerability to psychological distress in the COVID-19 scenario

In this section, we illustrate the methodological solution proposed in this article to address the considered problem; in particular, it is provided a high-level description of the two approaches that have been considered in this work to identify the potentially at-risk scenarios to be analyzed in terms of the psychological states, and in general the vulnerability to psychological distress of individuals with respect to the COVID-19 pandemic. In fact, as
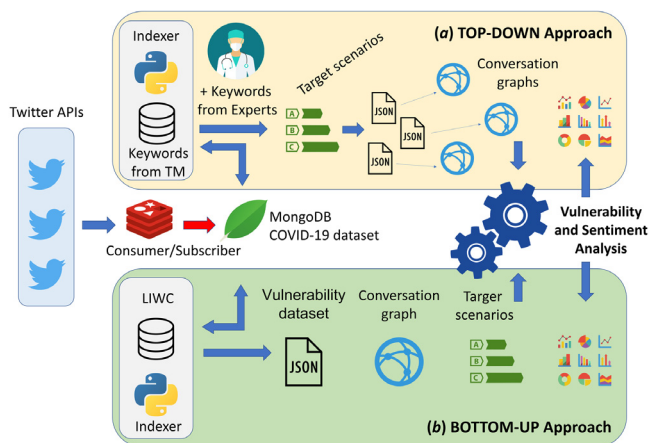
Fig. 1. The pipeline of the two approaches followed in this work.

outlined in the Introduction, we went down two paths: in the first direction, it was decided to follow a "top-down" strategy to select three specific at-risk scenarios that were estimated to be relevant by mental health specialists, later associated with discussion topics in Twitter and analyzed with respect to the presence of relevant lexicon identifiers of psychological vulnerability; in the second direction, it was decided to follow a solution that, given the tweets relating to vulnerable psychological states, proceeded in a "bottom-up" manner to identify the scenarios (also in this case associated with discussion topics) to be further investigated with respect to psychological distress vulnerability.

The pipeline encompassing both the top-down and the bottom-up approaches is illustrated in Fig. 1.

The top-down approach. In this approach, it was decided to study potential COVID-19 related lexicon identifiers of vulnerability to psychological distress with respect to three given at-risk scenarios (also denoted as *target* scenarios henceforth). Specifically, such scenarios have been identified by the mental health team, and concern:

(i) *social distancing* measures adopted to prevent the COVID-19 spread (in this case, the potential negative psychological effects could be related to loneliness, isolation, depression, etc.);

(ii) the debates on *vaccines* and *vaccination* campaigns (in this case, the potential negative psychological effects could involve fear, uncertainty, distrust, etc.);

(iii) *symptoms* felt by people and possible (or effective) *hospitalization* (in this case, the potential negative psychological effects could be anxiety, fear, panic, etc.).

In the top-down solution, given a dataset of tweets related to COVID-19 (details about the dataset used in this work will be provided in Section 4.1), three *topics* corresponding to the above-mentioned target scenarios around which conversations take place have been identified, where each topic is characterized by a set of associated *keywords*. Traditionally, the simplest way to identify topics within a textual document collection is to use unsupervised solutions such as *topic modeling*, which will be illustrated in detail in Section 3.1. In the top-down approach, to identify such topics and associated keywords, a mixed approach based on unsupervised learning and expert intervention has been employed.

Specifically, by following the structured pipeline illustrated in Fig. 1(*a*), we gathered COVID-19-related Twitter data, we applied to them topic modeling, we extracted a series of topics,

and, together with medical experts, we identified within them those that most easily referred to the three scenarios illustrated above. Once these topics were identified, a set of keywords provided by the experts were added to each topic to enrich it with domain-specific terms. At this point, the topics described by the associated keyword sets were used to re-filter the entire dataset and generate three sub-datasets, one for each topic considered. Subsequently, three specific *conversation graphs* were constructed on top of these datasets, whose contents related to the largest connected component of the resulting graphs were analyzed from the perspective of psychological vulnerability.

The bottom-up approach. In this second approach, instead of starting from potential target scenarios provided by domain experts, we considered all tweets characterized by the presence of keywords related to potential psychological distress, in order to identify only subsequently, in a bottom-up way, those that were the more interesting topics for our aims, and to associate them with target scenarios.

Specifically, according to the pipeline illustrated in Fig. 1(*b*), we considered the gathered COVID-19-related Twitter dataset and filtered it based on a depression-related lexicon that will be illustrated in Section 3.2.1. Through this filtering phase, we obtained a significant reduction in the number of total tweets, keeping, based on the lexicon used, only those tweets characterized by interesting psychological aspects. This has allowed the generation of a single conversation graph, this time related only to psychological distress; to the broader connected component of this graph topic modeling was applied, which allowed to identify three relevant topics in a bottom-up way. These topics, associated with target scenarios, were then analyzed with respect to the specificity of the vulnerability lexicon identifiers found.

3.1. Identification of keywords associated with target scenarios by means of topic modeling

In this section, we describe the *topic modeling* technique that was used in both the top-down and bottom-up approaches to extract topics (and associated keywords) from the data under consideration. In machine learning and natural language processing, a *topic model* is a type of statistical model for discovering the abstract topics that occur in a collection of documents, by capturing the hidden semantic structures in a text body. Intuitively, given that a document is about a particular topic, one would expect particular words to appear in the document more or less frequently. Hence, the topics produced by topic modeling techniques over a collection of textual documents are probability distributions over words that characterize each topic [33–35].

In both the top-down and bottom-up approaches presented in this paper, to perform topic modeling, *Latent Dirichlet Allocation* (LDA) [36] was applied. Specifically, it was carried out in distinct *runs*, by considering an increasing number of topics in each run (we remind that being topic modeling an unsupervised technique it is necessary to provide the number of topic in advance). In particular, a number of topics from 2 to 50 was taken into consideration. For each run (corresponding to a different number of topics), each returned topic is identified by an automatically-generated label and a list of associated keywords. The keywords are sorted by their frequency with respect to the considered topic.

Actually, there are various libraries and tools (including visual approaches) that can be employed to identify topics and help with analyzing them [37,38]. In this work, we proceeded by using the PyLDAvis tool,[2] which performs LDA and allows the identification, by means of an HTML interface, of the generated topics,

---
[2] https://pypi.org/project/pyLDAvis/

the associated keywords, their *saliency*, and *relevance* to the topic. In fact, the same word can appear in more than one topics, but it will have a different relevance with respect to different topics. In the employed approach, the saliency $s(w)$ of a term [37], and the relevance $r(w|T)$ of a term w for a topic T [38], are defined as follows:

$$s(w) = P(w) \cdot \left[\sum_t P(T|w) \cdot \log \frac{P(T|w)}{P(T)} \right] \qquad (1)$$

$$r(w|T) = \lambda \cdot P(w|T) + (1 - \lambda) \cdot \frac{P(w|T)}{P(w)} \qquad (2)$$

where the conditional probability $P(T|w)$ is the likelihood that the observed word w was generated by latent topic T, and the marginal probability $P(T)$ is the likelihood that any randomly-selected word w' was generated by topic T. Concerning relevance, λ is a weight parameter (where $0 \leq \lambda \leq 1$) determines the weight given to the probability of term w under topic T.

The use of this tool along with the *topic coherence* assessment [39], served in this paper to identify in a manual way (i.e., domain expert analysis) guided by an automatic method (i.e., topic coherence) the best extracted topics with respect to top-down and bottom-up approaches, as will be discussed in detail in the related sections.

3.2. Content analysis for assessing topic-dependent psychological vulnerability

The vulnerability analysis phase concerns, both for the top-down and the bottom-up approaches, the study of particular lexicon identifiers that can refer to psychological distress with respect to the different target scenarios identified through the two proposed approaches. In particular, we proceeded as follows:

(i) For each target scenario, we performed a two-fold *vulnerability analysis*, given the presence of words related to psychological vulnerability. In particular, (*a*) by identifying single keywords taken from a suitable depression-related lexicon that appeared more frequently within the different scenarios, and (*b*) by considering weighted terms having different importance with respect to the identification of psychological distress, as will be illustrated in more detail in the dedicated section;

(ii) For the content related to each distinct target scenario, *sentiment analysis* was performed, by considering both a lexicon-based and a semantic-based approach.

3.2.1. Vulnerability analysis

In the recent work described in [32], reference was made to the possibility of identifying features to measure depressive states in social media. In this work, in particular, we refer to features of a linguistic nature that can be extracted from the *Linguistic Inquiry and Word Count* (LIWC) lexicon [40]. As declared by its developers, the LIWC lexicon is the result of a transparent text analysis program that counts words in psychologically meaningful categories, allowing to capture the emotional states of individuals. In general, the spectrum of individual emotional states is very broad, and LIWC is quite general purpose. For this reason, in [32], to build a depressive dictionary, some sub-categories of LIWC have been employed, the ones more relative to depressive states, according to the authors. By considering these categories, and other features, the authors in their work design a depression marker model to build a list of words with a weight (value) attribute describing how depressive or non-depressive a word is.

Based on the same idea, but wanting to develop a more transparent approach that allows to maintain a clear link between

terms of LIWC and weighting of the same terms with respect to actually states of psychological vulnerability, in the current study we have implemented the following approach. First of all, the following categories of terms belonging to LIWC were identified by mental health experts as particularly significant with respect to the evaluation of vulnerability to psychological distress: *anger*, *anx*, *death*, *risk*, *sad*. Each category consists of a variable number of terms (ranging from a minimum of 103 terms for the category *risk*, to a maximum of 230 terms for the category *anger*). Categories may have some terms in common.

After the identification of these main categories, these terms were weighed, with respect to their importance in assessing psychological vulnerability. To do this, reference was made to the use of an external resource, that is to say a dataset, built on User-Generated Content, whose authors have declared to suffer from depressive symptoms. A similar solution was also proposed and adopted in [14], but using a different (and smaller) dataset for labeling than the one used in our work, and on a more limited number of LIWC categories. The dataset considered in this work has been made available in [41]. Specifically, the dataset has been generated in the context of the *early depression detection* problem, and the contents considered are posts published on the Reddit platform,[3] belonging to 137 users who have declared themselves depressed, for a total of 49,580 posts. In order to effectively weigh the terms present in LIWC with respect to this dataset, the frequency of the LIWC terms within the external resource was evaluated, considering in particular their normalized frequency with respect to the total number of terms present. Formally:

$$\omega_v(w) = \frac{f(w)}{N} \qquad (3)$$

where $\omega_v(w)$ denotes the *vulnerability weight* to be associated with the term w from the LIWC lexicon, $f(w)$ the frequency of the term w in the Reddit data collection, and N the total number of terms in the collection. In this weighting phase it was possible to verify how, out of a total of 1,315 terms present in LIWC, 724 terms were actually present in the dataset used for their weighting.

In order to verify, with respect to which LIWC category, which target scenarios were most at risk of psychological vulnerability, we also defined a *vulnerability score*, by multiplying the weight of the terms by their frequency in the scenario, and summing up these values together. Formally, given $w_1, w_2, \ldots, w_n$ the terms belonging to LIWC appearing in a given target scenario i, and $\omega_1, \omega_2, \ldots, \omega_n$ their associated weights, the vulnerability score $\sigma_v(i)$ for a target scenario has been computed as:

$$\sigma_v(i) = \sum_i \omega_i \cdot \frac{f(w_i)}{N} \qquad (4)$$

The results regarding the use of these vulnerability indicators will be presented in Section 5, concerning the psychological vulnerability analysis regarding the two approaches considered.

3.2.2. Sentiment analysis

Although not directly related to the concept of psychological vulnerability, it was decided to check, with respect to the content related to the target scenarios considered, their level of polarization. This may help confirm or not whether or not there may be a concordance between sentiment expressed by the posts and possible psychological suffering. Hence, to carry out *sentiment analysis*, we proceeded by considering two methodologies, one purely lexicon-based, and one semantic-based. The first is based on the use of the *Valence Aware Dictionary for sEntiment Reasoning* (VADER) [42], while the second is based on the use of

---

3 https://www.reddit.com/

the *Covid Twitter* BERT *model* (CT-BERT) [43]. The lexicon-based approach is pretty computationally inexpensive (an important aspect to be considered when dealing with huge amounts of data such as those that are disseminated on social media), and widely employed for general-purpose sentiment analysis [44]. The semantic-based approach has the advantage of being able to consider semantic and contextual aspects with respect to the domain of interest considered, although it is computationally more onerous.

Concerning VADER, it employs a human-generated English sentiment lexicon, where lexical features (i.e., words) are labeled according to their semantic orientation as *positive*, *negative*, or *neutral*, also expressing the sentiment intensity for each lexical feature. In addition, VADER manages a proper handling of punctuation (for example, the exclamation point (!) increases the magnitude of the intensity without modifying the semantic orientation, e.g., *Coronavirus won't stop us!*), as well as capitalization (ALL-CAPS is used to emphasize a sentiment-relevant word in the presence of other non-capitalized words, e.g., *Coronavirus WE WILL WIN THIS TOGETHER*), degree modifiers (which impact on sentiment by either increasing or decreasing its intensity, e.g., *I am extremely worried about Coronavirus*), sentiment-laden slang words, and emoticons/emoji. By means of VADER, the assessment of tweets polarity was performed by computing the so-called polarity *compound score* for each tweet. The compound score is computed as the sum of all lexicon ratings associated with words, normalized between -1 (extremely negative) and $+1$ (extremely positive). Based on the obtained compound scores, we computed the proportion of the resulting negative (compound score $\leq$ -0.05), neutral (compound score between -0.05 and 0.05), and positive (compound score ≥ 0.05) tweets for the considered time period.

To perform semantic-based sentiment analysis, we considered the model described in [45]. Such model is pre-trained on a corpus of COVID-19-related tweets. To perform sentiment analysis, the original CT-BERT model [43] – trained on 22.5 million tweets collected between January and April, 2020 containing at least one of the keywords "wuhan", "ncov", "coronavirus", "covid", or "sars-cov-2" – has been adapted by adding a so-called fine-tuning layer (i.e., a single neural layer) to the model, trained on the SemEval-2017 Task 4 (Sentiment Analysis in Twitter) dataset [46]. This made it possible to disambiguate, in the specific COVID-19 context, the neutrality of tweets originally identified by VADER as positive (in most of the cases) or negative, as detailed in Section 5.

4. Instantiation of the top-down and bottom-up approaches for target scenarios construction

In this section, we illustrate the instantiation of the proposed top-down and bottom-up approaches with respect to a dataset of tweets related to COVID-19 that was collected for the occasion. This dataset is described in Section 4.1, while Sections 4.2 and 4.3 describe, respectively, the steps performed to identify target scenarios with respect to the top-down approach and those performed with respect to the bottom-up approach.

4.1. A COVID-19 dataset for vulnerability analysis on the twitter microblogging platform

As illustrated in Section 2, there are numerous studies that especially in the last year have focused on the use of Twitter to make health-related analyses relating to COVID-19. This is also favored by the fact that, with the spread the COVID-19 pandemic, and to help scientists from different disciplines to evaluate its effects on society, Twitter has made available, on request, the so-called COVID-19 *stream endpoint*, through which it is possible to

collect, by means of the platform's API, a considerable amount of tweets related to different keywords that identify the pandemic.[4] Through access to this data stream, we gathered, from 15 August until 31 December 2020, around 262 million tweets.

However, the purpose of our work is not to focus on the study of any possible discussions related to COVID-19 in general; instead, we focus on the identification and analysis of specific lexicon identifiers of mental health-related vulnerability that is probably favored by specific aspects of the pandemic, and the way people have lived in this long period of isolation, fear, stress, and so on. For this reason, the original dataset has been filtered on the basis of certain topics and certain psychological states identifiers, as introduced in Section 3, and as detailed in the next two sections with respect to the two considered approaches.

4.2. Instantiation of the top-down approach

The instantiation of the top-down approach with respect to the COVID-19 dataset under consideration concerned: (*i*) the application of topic modeling to extract topics and keywords from the original dataset and to allow the experts to select the most significant topics and keywords with respect to the initially established target scenarios; (*ii*) the enrichment of the automatically extracted keywords with keywords specific to each scenario and provided by the domain experts; (*iii*), the filtering of the original dataset based on such lists of words associated with each scenario, in order to identify the content related to these scenarios and generate related conversational graphs.

4.2.1. Topic modeling to identify top-down target scenarios

As illustrated in Section 3.1, multiple runs have been performed using LDA, for a number of topics ranging from 2 to 50.[5] To provide an idea, Table 1 show the identified topics and the respective associated keywords (the top-30 terms) in the case of 3 and 7 topics considered.

As it can be seen from the table, as the number of topics that we try to identify in the collection of documents increases, the more these topics are able to be discriminating. For example, by forcing the generation of only 3 topics, we can say little about their characteristics; topic 1 is constituted by mixed terms referring to distinct aspects related do COVID-19; together with topic 2, it is, in any case, the most health-related (terms in magenta). In fact, topic 3 seems to be more related to politics (terms in green), which in the top-down approach has not been deemed as an interesting at-risk scenario according to the considered psychological issues.

With a number of topics equal to 7, we can see how topics start to be more differentiated; in fact, in addition to the political topic that in this categorization appears more identified by topic 7 (terms in green), we can see how in topics 5 and 6 words appear that are related to (Christmas) holidays, to spending time with friends and relatives, and to a sense of hope (terms in pink). We can also see how some of the topics contain keywords that are already pretty closer to the target scenarios that we are interested in analyzing in this work. For example, topic 2 seems to be most correlated to hospitalization (terms in purple), while in topics

---

[4] The COVID-19 stream endpoint return tweets based on the Twitter's COVID-19 Tweet annotation, and a set of defined parameters giving a comprehensive view of the conversation around this topic. For more information: https://developer.twitter.com/en/docs/labs/covid19-stream/overview

[5] In the top-down approach, topic modeling has been applied, for efficiency issues, to 10% of the tweets collected every day, being in fact just a phase of support to the domain experts for the identification of a set of potential keywords considered significant for the next filtering phase, an activity anyway to be carried out manually in such a top-down strategy.

Table 1

Examples of topic modeling with LDA applied to the random dataset considered for a number of topics equal to $k = 3$ and $k = 7$.

Topic	Keywords
Run $k = 3$	
1	Lockdown, relief, government, health, public, time, christmas, country, workers, vaccines, state, money, congress, world, work, week, stimulus, lockdowns, countries, news, americans, restrictions, read, vaccination, millions, support, immunity, december, deaths, businesses.
2	Cases, deaths, year, home, test, vaccines, christmas, pfizer, quarantine, time, work, hospital, spread, million, health, care, patients, death, safe, family, live, stay, rate, world, life, house, died, number, positive, tests.
3	Trump, wear, wearing, tested, positive, realdonaldtrump, president, biden, americans, days, good, died, china, year, election, lives, social, time, media, line, america, fucking, sick, stop, news, country, lost, remember, life, family.
Run $k = 7$	
1	Lockdown, spread, government, lockdowns, work, world, strain, immunity, stop, rate, time, open, vaccines, restrictions, south, herd, great, year, support, variant, december, countries, state, asymptomatic, schools, vaccination, country, week, govt, read.
2	Cases, deaths, health, death, million, public, number, lives, total, high, state, record, states, news, numbers, reported, risk, world, daily, hospitals, hospital, crisis, data, country, canada, breaking, days, united, report, patients.
3	Trump, vaccines, pfizer, realdonaldtrump, biden, china, president, america, fauci, moderna, news, american, doses, americans, election, world, years, country, media, administration, time, force, money, hoax, truth, year, history, effective, medical, science.
4	Positive, wear, test, tested, wearing, social, distancing, vaccinated, face, negative, tests, home, good, twitter, hands, christmas, life, year, symptoms, days, tweet, holiday, phone, quarantine, family, night, post, game, time, season.
5	Lockdown, christmas, died, year, fucking, days, government, family, lost, americans, hospital, patients, months, care, workers, money, million, weeks, holidays, homes, politicians, child, business, week, finally, country, feel, anti, friends, billion.
6	Stay, time, wear, safe, home, love, quarantine, year, vote, hope, good, wearing, literally, life, line, happy, family, person, worry, sick, middle, black, shut, country, girls, healthy, merry, voting, face, christmas.
7	Relief, congress, work, americans, care, health, workers, stimulus, house, senate, republicans, trump, checks, pass, working, package, mcconnell, government, mitch, millions, white, healthcare, home, american, court, staff, plan, president, democrats, tested.

1, 3, and 4 appear some terms related both to social distancing (terms in orange) and to vaccines (terms in cerulean).

However, it is clear that this does not represent in our eyes a clustering that is still totally effective to identify the target scenarios. Hence, having the set of keywords associated with tentative topics in different runs, it was necessary to carry out a manual analysis to better identify the topic model parameters that best represented the three target scenarios we wanted to study. This manual analysis was performed both by means of the help provided by the PyLDAvis tool introduced in Section 3.1, and by referring to the topic coherence values calculated as described in [39] and illustrated in Fig. 2.

As can be seen from the figure, the highest values of topic coherence occur for a number of topics equal to 9, 20, 38, and 48. Fig. 3 illustrates the use of the tool for a number of topics equal to 9, from which the experts have extracted the significant keywords related to the three target scenarios considered.
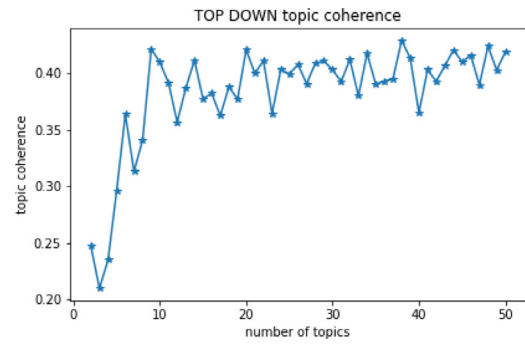

Fig. 2. Topic coherence values over the 50 runs performed in the top-down approach.

Table 2

Keywords extracted through the application of topic modeling with 9 topics and association with the respective target scenarios.

Target	Keywords
Social distancing	Distance, distancing, family, holiday, home, lockdown(s), mask, quarantine, restrictions, safe, smartworking, social, spread, stay, wear, wearing.
Vaccines & vaccinations	Asymptomatic, billion, biotech, drug(s), fear, government, herd, immunity, money, negative, paid, pfizer, positive, rich, test(s), tested, testing, trial, vaccination(s), vaccine(s).
Symptoms & hospitalization	Bed(s), cancer, care, doctor(s), hospital(s), nurse(s), patient(s), room, strain, symptom(s), treatment(s).

Such keywords, which have been extracted in particular from topics with labels 3, 6, and 9, are illustrated in Table 2. In particular, with respect to the keywords associated with the extracted topics, a *keyword selection* phase was performed. This was obtained both automatically, by removing overlapping words among topics, by keeping the one with higher saliency and relevance with respect to the topic, and thanks to the team expertise, to remove those keywords deemed unsuitable or not particularly significant with respect to the target scenario.

Furthermore, other domain-specific keywords relating to the three situations of interest were also provided, as illustrated in the next section.

4.2.2. Enrichment of target scenarios by means of keywords provided by experts

The unsupervised automatic method for topic identification contributed to extract a first useful list of keywords to be associated with the target scenarios of interest. However, we decided to consider other potentially relevant *domain-specific keywords* provided by medical experts. Therefore, in addition to those illustrated in Table 2, the following terms have been added to each target scenario. Specifically, with respect to *social distancing*, they are:

$$\mid mask - facemask - gloves - hand\ (hygiene) \mid$$

The above-mentioned terms have been extracted from the *COVID-19 table of PPE with description and related standard (simplified version),*[6] by the *World Health Organization* (WHO).[7] With
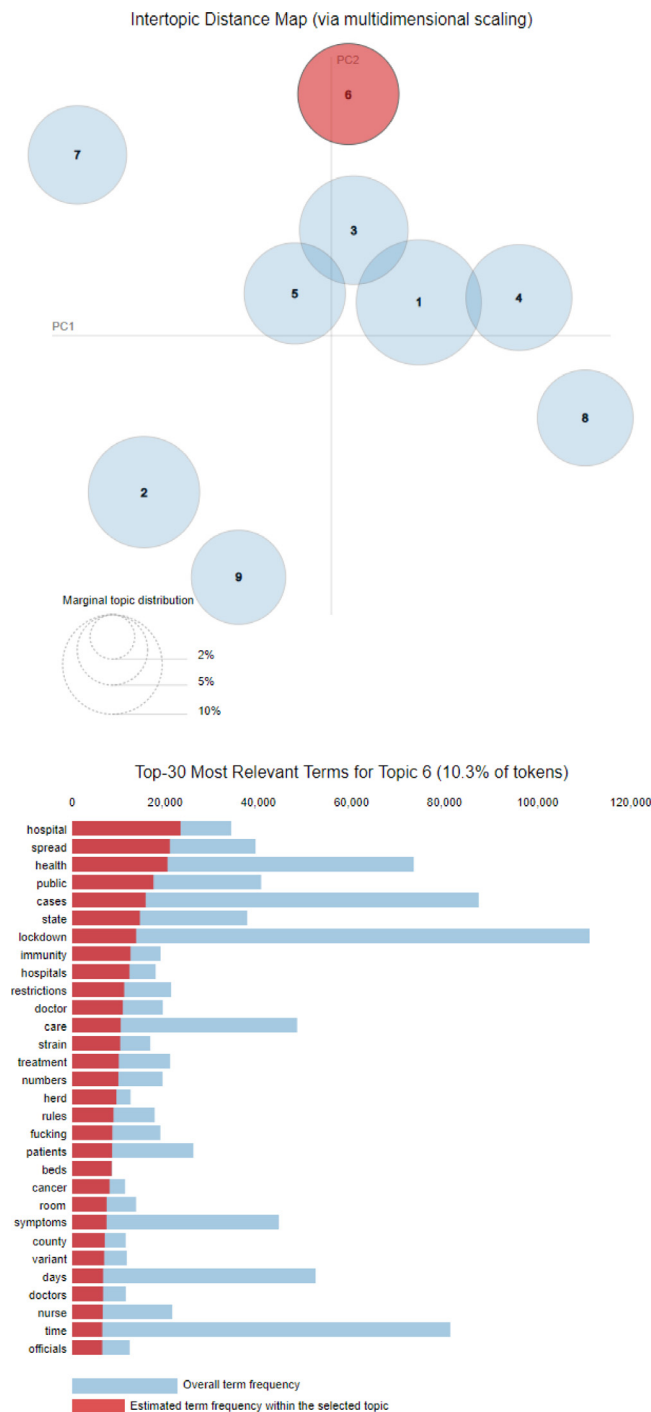
Fig. 3. Interface of the tool used to interpret the topics extracted by choosing a variable number of topics. In particular, in the figure, the number of extracted topics is equal to 9, and the selected topic is the one labeled as 6.

respect to *vaccines & vaccinations*, the following keywords, extracted by the WHO's *Draft landscape of COVID-19 candidate vaccines*,[8] have been added:

> *DNA – RNA – replicating – inactivated – vector attenuated – recombinant – particle – protein – subunit moderna – astrazeneca*

---
[8] https://www.who.int/publications/m/item/draft-landscape-of-covid-19-candidate-vaccines

Finally, with respect to *symptoms & hospitalization*, the following keywords, extracted by the list of symptoms provided by the *Centers for Disease Control and Prevention*,[9] have been added:

> *fever – chills – breath (*and/or *breathing) – fatigue (muscle* or *body) ache – headache – taste – smell (sore) throat – congestion – runny (nose) – nausea vomiting – diarrhea – antibody (*or *antibodies) – test serology – pharyngeal (*and/or *pharynx) – swab molecular – immunological – antigen*

4.2.3. Filtering the dataset by means of keywords identifying target scenarios

Based on the keywords extracted from the topic modeling phase and those added by the experts, the three resulting sets of keywords representing the three target scenarios were used to filter the original COVID-19 dataset, which, we recall, consisted of around 262 millions of tweets. In this way, it was possible to generate three sub-datasets that contain only those tweets that refer to the target scenarios of interest. The characteristics of these datasets, and of the resulting conversation graphs, are illustrated in Table 3.[10]

In particular, the table shows the number of filtered tweets (# tweets) associated with each target scenario, and the characteristics of the resulting conversation graphs. The graphs have been built by considering users whose tweets contain at least one term belonging to the related target scenario, and edges among users have been built by considering *retweet*, *mention*, and *quote* actions. The resulting structure was a weighted graph, where weights on edges represent the number of interactions among users. In the table, the number of nodes (# nodes) and edges (# edges) relating to the conversation graphs built this way are indicates, as well as the degree of each graph (*G* degree), and the number, for the largest connected component (CC) of each graph, of nodes (CC nodes) and edges (CC edges), together with its degree (CC degree).

Furthermore, the data relating to the largest connected component when the edges with weight equal to 1 were removed are also provided, in particular its number of nodes (CC2 nodes), edges (CC2 edges), and its degree (CC2 degree). Only the data relating to this latter connected component were considered in our subsequent analyses, as only the nodes that have had at least two interactions between them have been taken into consideration. This choice was dictated by the desire to analyze the most closely interacting communities, bearing in mind the concept, illustrated in the Introduction, of possible emotional contagion discussed in the literature [23].[11]

4.3. Instantiation of the bottom-up approach

The instantiation of the bottom-down approach with respect to the COVID-19 dataset under consideration concerned: (*i*) the filtering of the dataset based on the depression-related terms that were extracted through the approach described in Section 3.2.1 and the construction of the related (single) conversational graph, and (*ii*) the application of topic modeling to the content extracted from that conversational graph to identify the most significant topics to be associated with target scenarios on which to assess psychological vulnerability.

---
[9] https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html

[10] In the table, the number of tweets may be lower than the number of users because only tweets referencing COVID-19 were made available through the Twitter stream endpoint and considered in this work; in addition, a single tweet related to COVID-19 may be linked to more than one user if a response to the tweet is provided or if it is retweeted or mentioned.

[11] It should also be noted that in this work we are not interested in analyzing the behavior of possible outliers, and that the members of the largest connected component still constitute about 80% of the total virtual community considered.

Table 3

Characteristics of the datasets and conversation graphs associated with the three target scenarios after the filtering phase on the basis of the keywords extracted by means of topic modeling and the intervention of experts.

Datasets	Social distancing	Vaccines & vaccinations	Symptoms & hospitalization
# tweets	3,347,407	944,108	632,200
# nodes	3,900,140	3,518,257	1,047,215
# edges	8,398,646	867,068	1,922,023
G degree	4.3068	4.4893	3.6707
CC nodes	3,509,001	1,445,218	909,089
CC edges	8,146,111	3,441,020	1,832,946
CC degree	4.6430	4.7619	4.0325
CC2 nodes	212,012	102,497	49,685
CC2 edges	390,702	183,200	76,012
CC2 degree	3.686	3.575	3.060

Table 4

Characteristics of the *vulnerability dataset* and the associated conversation graph.

	Vulnerability dataset
# tweets	1,000,000
# nodes	677,286
# edges	1,013,258
Graph degree	2.9921
CC nodes	579,469
CC edges	949,646
CC degree	3.2776
CC2 nodes	221,498
CC2 edges	591,675
CC2 degree	5.3425

Fig. 4. Topic coherence values over the 50 runs performed in the bottom-up approach.

4.3.1. Filtering the dataset based on depression-related terms

Through this filtering phase, we obtained the number of tweets illustrated in Table 4, which allowed the generation of a single conversation graph, this time only related to psychological distress, with the characteristics there illustrated.

On the basis of this set of tweets, a topic modeling activity was carried out on those belonging to the largest connected component of the graph where edges with weights equal to 1 were excluded (CC2), to evaluate if it was possible to recognize topics referable to scenarios of interest, which in the top-down case had been indicated directly according to the expertise of the mental health team.

4.3.2. Topic modeling to identify bottom-up target scenarios

As in the case of the top-down approach, the same topic modeling technique illustrated in Section 3.1 has been applied to the bottom-up approach, performing the same number of runs from 2 to 50.[12] In this case, the values of topic coherence are shown in Fig. 4 and appear higher for a number of extracted topics of 32, 36, 40, and 42.

Again, thanks to such topic coherence values and the use of the visual tool to analyze the extracted topics, the medical expert team evaluated as suitable to identify possible target scenarios the keywords associated with labels 1, 12, and 27, when the number of considered topics was 40. Despite the high number of generated topics, it was observed that there were always three or four topics of greatest importance to the aim of our work, which were simply divided into several sub-topics that were semantically overlapping when the number of topics extracted increased.

The list of the 15 most frequent words for each considered topic is illustrated in Fig. 5. Topic 1 can quite clearly represent the scenario concerning the adoption of *social distancing* measures and the use of *protective devices* and *hygiene measures*: in fact, we

---

[12] In the bottom-up approach, given its different nature w.r.t. the top-down solution and the significant reduction in the size of the dataset considered due to the preliminary filtering phase, the topic modeling activity was applied to the entire dataset.
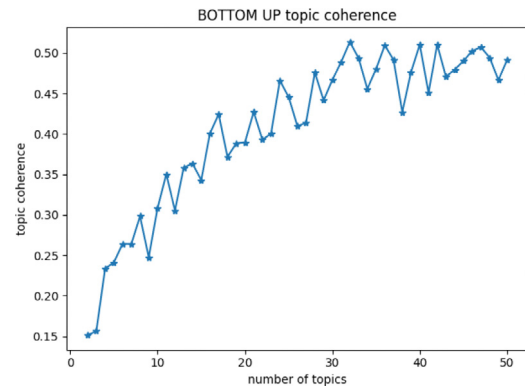
found keywords related to these aspects, e.g., wear(ing), mask(s), social, distancing, home, stay, safe, etc. Topic 12 was mostly related to *tests* for the detection of positivity to the virus and *hospitalization*. Finally, topic 27 was mainly related to *politics*.

From this topic modeling activity it emerged how, compared to the topics (and, hence, the target scenarios) initially considered in the top-down approach, there are actually some overlaps, and how the political sphere, initially not taken into consideration, is somehow connected to the presence of relevant lexicon identifiers of psychological vulnerability. As for the vaccines topic, expressly taken into consideration in the top-down approach, this emerged through the presence of some keywords associated with different cross-cutting topics, especially as the number of considered topics increases, but it did not emerge as an independent topic (and therefore as a target scenario). This can perhaps be traced back to the particular time period in which the tweets were collected, in which the vaccine development stage was still very embryonic.

5. Results of the vulnerability analysis for the top-down and bottom-up approaches

The top-down and bottom-up approaches, instantiated on the dataset of tweets related to COVID-19, both ultimately resulted in a set of content filtered based on particular keywords related to three target scenarios. In this section, these contents are analyzed with respect to the presence of lexicon identifiers of psychological vulnerability and to polarity, based on the measures and approaches described in Section 3.2.

5.1. Vulnerability analysis for the top-down approach

Both on the basis of the non-weighted words present in LIWC and on the basis of the weighted dictionary generated by the proposed approach, a double vulnerability analysis was performed. A
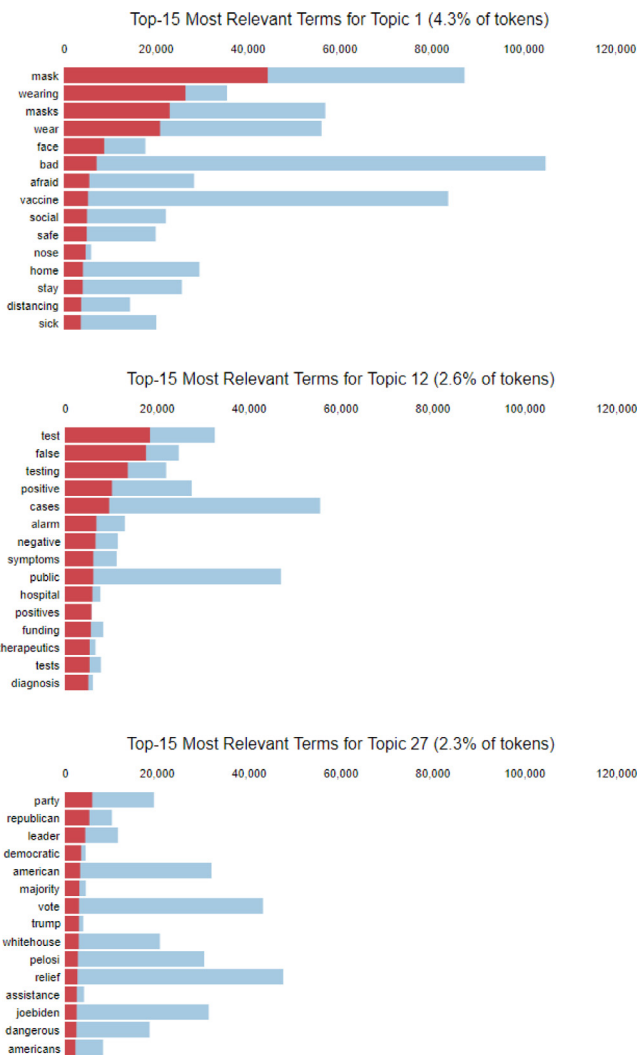
Fig. 5. Top-15 terms for the topics 1, 12, and 27, with a number of extracted topics equal to 40.

Table 5
Global vulnerability scores for the three (top-down) target scenarios.

Target scenarios		
Social distancing	Vaccines & vaccinations	Symptoms & hospitalization
0.0038	0.0026	0.0037

first analysis was made, for each target scenario, of the most top-5 frequent words belonging to the LIWC categories *anger*, *anx*, *risk*, *sad*, as illustrated in Figs. 6–8.[13]

In addition, Table 5 summarizes the global vulnerability scores computed for each target scenario, while Fig. 9 illustrates the proportion of the vulnerability scores computed for each category of LIWC terms with respect to each scenario.

Finally, Table 6 illustrates the trend of the sentiment for the three target scenarios using the VADER lexicon-based approach, while Table 7 presents the results related to the usage of the semantic-based CT-BERT model.

---

[13] The figure relating to the *death* category was not provided as the terms used and their distribution did not present substantial differences among target scenarios.
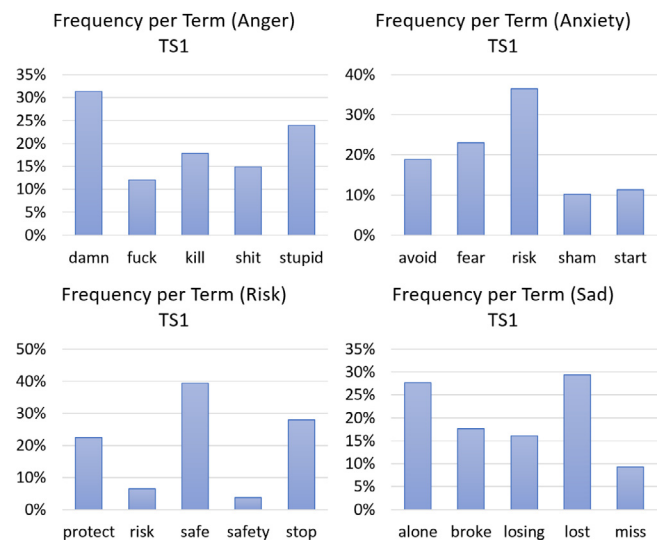

Fig. 6. Frequencies of the top-5 terms that appear in the target scenario *social distancing* with respect to the four LIWC categories.
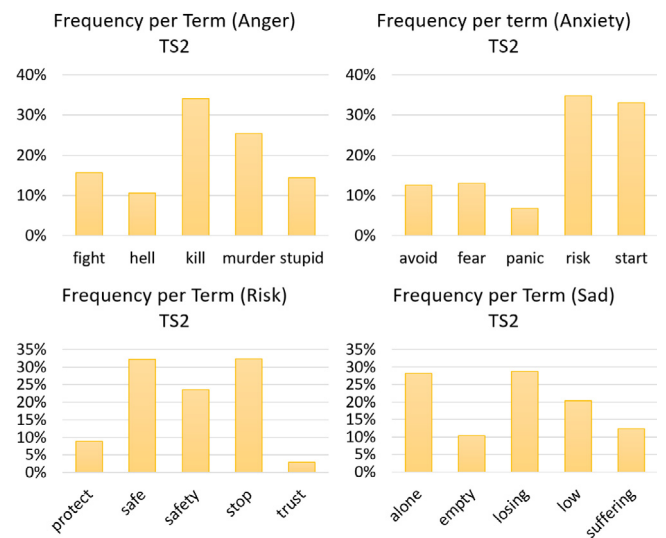

Fig. 7. Frequencies of the top-5 terms that appear in the target scenario *vaccines & vaccinations* with respect to the four LIWC categories.

5.2. Vulnerability analysis for the bottom-up approach

Using the same vulnerability analysis methodology detailed for the top-down approach, also in this case we provide the results of the presence of terms belonging to LIWC, and of the presence of terms that are more significant as identifiers for indicating psychological vulnerability.

First of all, the top-5 most frequent LIWC terms that appear in the considered scenarios are illustrated in Figs. 10–12, with respect to the distinct LIWC categories illustrated in the case of the top-down approach.

Secondly, the vulnerability values associated with the three (bottom-up) target scenarios are indicated in Table 8.

The vulnerability values are much lower than those obtained in the top-down approach because the number of terms on which these values have been normalized is much higher in the case of the bottom-up approach, which started from a much larger dataset of tweets. Finally, in Fig. 13, the proportions of the vulnerability scores for the three target scenarios with respect to the individual LIWC categories are presented.

Table 6

Sentiment analysis values obtained by VADER on the three considered (top-down) target scenarios.

Target	Social distancing	Vaccines & vaccinations	Symptoms & hospitalization
Positive	39%	46%	40%
Neutral	20%	15%	12%
Negative	41%	39%	48%

Table 7

Sentiment analysis values obtained by employing CT-BERT on the three considered (top-down) target scenarios.

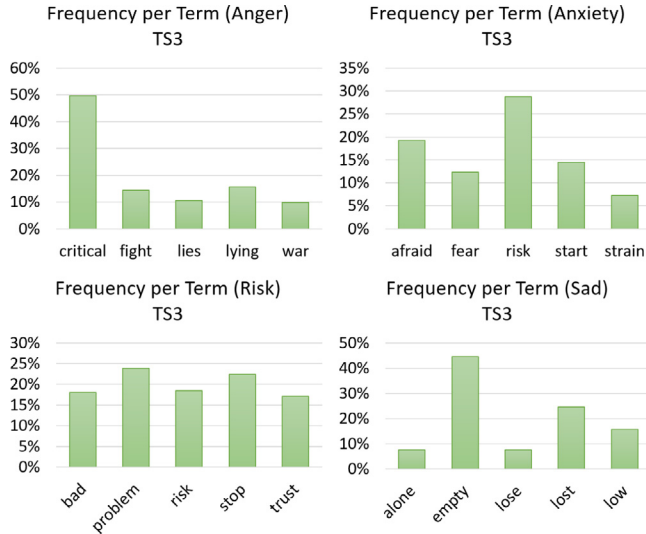| Target	Social distancing	Vaccines & vaccinations	Symptoms & hospitalization
Positive	10%	8%	4%
Neutral	40%	35%	46%
Negative	60%	57%	50%

Fig. 8. Frequencies of the top-5 terms that appear in the target scenario *vaccines & vaccinations* with respect to the four LIWC categories.
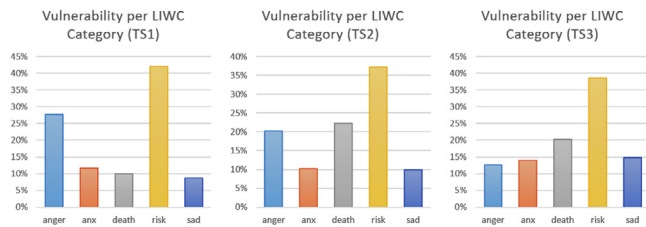

Fig. 9. Vulnerability scores per category with respect to the three (top-down) target scenarios.

Table 8

Global vulnerability scores for the three (bottom-up) target scenarios.

Target scenarios		
Social distancing & protection	*Tests & hospitalization*	*Politics*
0.0011	0.0007	0.0005

With respect to the *polarity* associated with tweets related to the considered target scenarios, by using the same sentiment analysis techniques illustrated in Section 3.2.2, results are illustrated in Table 9 (VADER) and in Table 10 (CT-BERT).

5.3. Discussion

By referring to the results obtained and illustrated in the previous sections, allowing to capture the potential presence
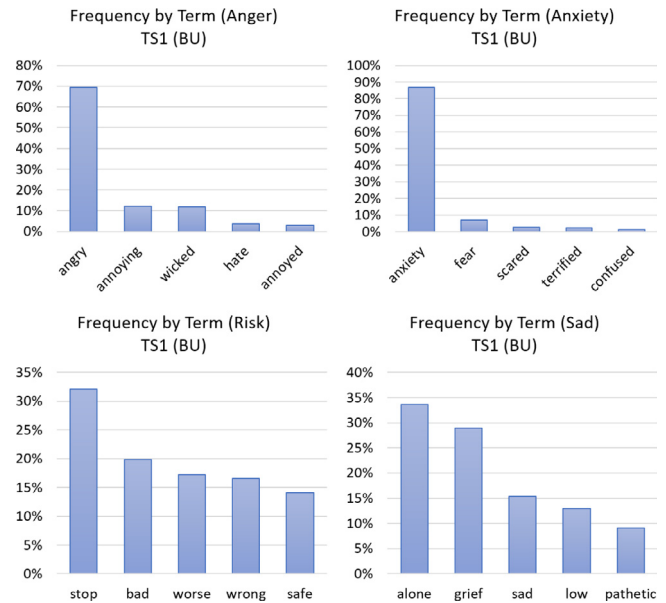

Fig. 10. Frequencies of the top-5 terms that appear in the (bottom-up) target scenario *social distancing and protection* with respect to the four LIWC categories.
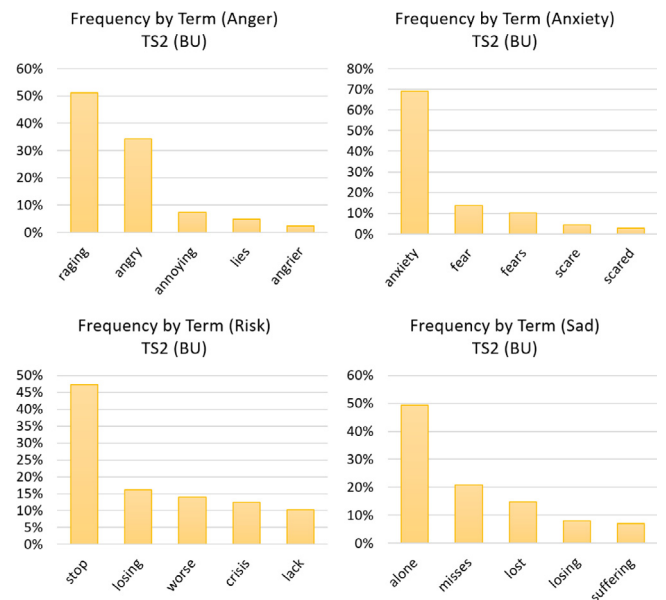

Fig. 11. Frequencies of the top-5 terms that appear in the (bottom-up) target scenario *tests & hospitalization* with respect to the four LIWC categories.

Table 9

Sentiment analysis values obtained by VADER on the three considered (bottom-up) target scenarios.

Target	Social distancing and protection	Tests & hospitalization	Politics
Positive	20%	7%	64%
Neutral	2%	2%	2%
Negative	78%	89%	34%

Table 10

Sentiment analysis values obtained by employing CT-BERT on the three considered (bottom-up) target scenarios.

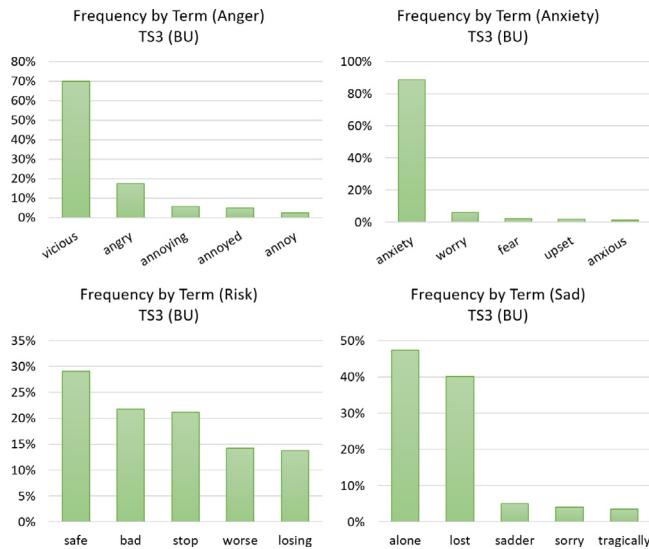| Target	Social distancing and protection	Tests & hospitalization	Politics
Positive	8%	7%	9%
Neutral	23%	23%	26%
Negative	69%	70%	65%

Fig. 12. Frequencies of the top-5 terms that appear in the (bottom-up) target scenario *politics* with respect to the four LIWC categories.
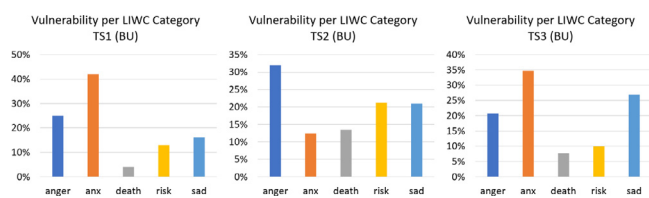

Fig. 13. Vulnerability scores per category with respect to the three (bottom-up) target scenarios.

of vulnerability to psychological distress in the context of the conversations held on Twitter with respect to the COVID-19 pandemic, some considerations about the different scenarios and considered approaches are here provided. Given the hybrid, multidisciplinary nature of the proposed approach, the results can be interpreted qualitatively, by focusing on the experience of the medical staff.

5.3.1. Main findings

The first observation that emerges is that, considering both the top-down and bottom-up approaches, we have been able to capture target scenarios that are at least partially overlapping, apart from the case of vaccines and vaccinations that may have been influenced by the particular period of data collection. This is particularly evident in the bottom-up approach, which is totally data-driven. However, there are sometimes significant differences

between the two approaches in the vulnerability and content polarity analyses that merit further discussion.

Vulnerability analysis and top-down approach. From the obtained results, we can observe how the most frequent terms regarding the four LIWC categories illustrated in Figs. 6–8 change, at least in part, depending on the considered scenario. If we refer to the *social distancing* scenario, we can see for example that, compared to the *anger* category, there are more terms that refer to annoyance and protest, while the same category relating to the *vaccines & vaccinations* scenario has a couple of terms, i.e., *kill* and *murder*, which may perhaps refer to the concern of serious contraindications and reactions. Regarding the category of *anxiety*, it seems that the target scenarios *social distancing* and *symptoms & hospitalization* involve more aspects related to risk, worry, fear. The concept of *risk* emerges in all three scenarios, even if words related to remaining *safe* and to *safety* appear more connected to *social distancing* and *vaccines & vaccinations*.

Concerning the global vulnerability scores illustrated in Fig. 9, we can observe that the *social distancing* and *symptoms & hospitalization* scenarios seem to be more at risk. If we then analyze in detail the LIWC categories that have higher vulnerability scores, in Fig. 9, we realize that the lexicon identifiers that indicate greater psychological vulnerability fall into the *anger* and *risk* categories in particular for the target scenarios *social distancing* and *vaccines & vaccinations*. Compared to the first scenario, the second has a large increase in vulnerability identifiers that fall into the *death* category, an aspect that is shared by the *symptoms & hospitalization* target scenario. Unlike the first two, however, the third scenario shows a limited number of vulnerability identifiers in the *anger* category. The lexicon identifiers belonging to the *anx* and *sad* categories are roughly in the same proportion in all three scenarios.

Vulnerability analysis and bottom-up approach. In this case we note, in Figs. 10–12, that the *anger* category terms are very similar for *social distancing & protection* and *tests & hospitalization*, while there is an interesting spike of the term *vicious* in relation to the *politics* scenario. If we consider the *anx* category, the terms are pretty similar for the three scenarios; as regards the *risk* category, the terms are quite similar for the first two scenarios, while for the *politics* scenario there is a peak of the term *safe*, perhaps a symptom of the need for protection that is expected from politics. The sense of loneliness emerges overwhelmingly with respect to all three scenarios in the *sad* category, accompanied by a peak of the term *lost* in the political scenario. In association with the consideration made above, it could refer to the need for security which is not currently met.

Also in this case, referring to global vulnerability scores in Table 8 (the reason for such low values has been explained in the dedicated section) we can observe how the most at-risk scenarios (with respect to psychological vulnerability) seem to be the first

two. Finally, if we refer to Fig. 13, which reports the percentages of the vulnerability scores referred to the various LIWC categories, we can note that *anxiety* seems to be the emotional state that most distinguishes the *social distancing & protection* scenario; *anger*, *sadness* and emotional states linked to a possible *risk* characterize the *tests & hospitalization scenario*; *anxiety* and *sadness* are the emotional states most linked to the *politics* scenario.

Comparison between the two approaches. From both approaches, it emerges that the scenarios *social distancing* and *symptoms & hospitalization* from the top-down approach, and the almost corresponding scenarios *social distancing & protection* and *tests & hospitalization* from the bottom-up approach, are those that are particularly affected by potential psychological vulnerability. Thus, in this case, both approaches lead to a similar result.

On the contrary, it was possible to identify, by means of the bottom-up approach, that even when users talk about topics in relation to *politics* and COVID-19, they use as well important lexical identifiers related to potential psychological vulnerability, an aspect that was not initially considered in the top-down approach. In addition, finding many important lexicon identifiers in the *anger* category in the bottom-up approach, somewhat contradicts what was evidenced by the top-down one, which had demonstrated a low anger score for the *symptoms & hospitalization* scenario. This may be due to the fact that in the second case, more than the symptoms, reference was made to the test phase with respect to COVID-19 positivity, which emerged more than the symptom aspect in association with hospitalization. This could mean that the two scenarios are actually interpreted very differently and actually represent two different situations: the fear of being hospitalized having some symptoms, and possible dissatisfactions related to the screening procedures. However, these and other hypotheses certainly deserve further analysis in order to be fully confirmed.

Sentiment analysis. Discussions related to sentiment analysis computed w.r.t the two approaches considered, depend heavily on whether VADER or CT-BERT is used to perform this task. For both top-down and bottom-up approaches, using CT-BERT to analyze content polarity allows us to "correct" some distortions caused by using the lexicon-based approach.

For example, when considering VADER, we note how the conversations about *social distancing* are more or less equally distributed between a positive and negative sentiment in the case of the top-down approach, while they are more clearly negative in the bottom-up approach with regard to the target scenario *social distancing and protection*. However, when we refer to CT-BERT, we see how sentiment is basically trending negative for all three scenarios in both approaches.

As for the two non-overlapping target scenarios in the two approaches, namely *vaccines & vaccinations* (top-down) and *politics* (bottom-up), they both show a mostly positive sentiment when considering VADER, while they turn to a negative sentiment as the other scenarios when considering CT-BERT.

5.3.2. Limitations

The proposed work, in part because of its interdisciplinary nature, the need to largely use medical experience through the definition of hybrid solutions, and the novelty of the problem, meant that choices had to be made and results had to be analyzed in ways that can be considered, in some aspects, limitations of the approach.

For example, the current work is based on Twitter as a widely adopted social media. However, Twitter is extremely popular in countries like the US and Japan, but much less in countries like Russia, China, and some EU countries. Findings might be influenced by the different popularity and usage of Twitter in

various countries, against other popular social media platforms (e.g., Facebook, Reddit, SnapChat, Instragram, Glitch). Thus, further research should consider potential social media competitors.

Regarding the technical solutions proposed and the analysis of the results, we are aware that the former do not constitute an absolute novelty, but we believe that they are if used in the way we did within an innovative field such as the study of mental well-being through social content analysis; as for the results obtained, they are essentially a qualitative assessment of the proposed approach, which was interpreted by the domain experts based on their experience in the field.

At present, the available literature does not yet provide adequate, publicly accessible resources for a strictly comparative evaluation. However, we have relied on some shared aspects and premises in the development of the current work.

6. Conclusions

Social media are nowadays the tool through which people can freely publish content that refers to various aspects of their personal life, and therefore also to their state of mental health and psychological vulnerability.

Over the past year, we have all had to deal with the effects the COVID-19 pandemic has had on our lives. For many people, this has been particularly difficult from a psychological point of view. The idea from which this article was born is that many individuals will have used social platforms, Twitter in particular, to talk about their discomfort, which, in our opinion, can concern various areas related to the pandemic. One can feel anxious about fearing contagion, depressed about being isolated and distant from other people, wary and afraid of the possibility of being vaccinated and hospitalized.

In our opinion, it was therefore necessary, starting from content disseminated on Twitter relating to the pandemic, to better identify these situations and the degree of psychological vulnerability of individuals with respect to them. This problem, in order to be addressed, undoubtedly requires the participation of experts in the psychological/psychiatric field and the use of information technologies for the analysis of large amounts of data. This led to the study of the problem from two points of view, one top-down, based on the definition by experts of target scenarios, and one bottom-up, consisting above all of a data-driven strategy. Through both these approaches, it was possible to verify how interesting outcomes can emerge, useful to mental health experts in the face of the use of social media, especially in situations as extreme as a pandemic can be.

Multiple future developments can be considered. In this work we have considered some techniques (e.g., topic modeling and sentiment analysis) taken from the literature, albeit employed and parameterized according to the specific scenario. From this perspective, a comparison with other techniques can certainly be of interest, also as regards the development of more semantically relevant approaches for the weighting of the dictionary of terms with respect to their importance in verifying the state of psychological vulnerability. Furthermore, it could be useful to deepen the investigation of content-aware community detection algorithms to capture more relevant and specific aspects of the emotional contagion.

Finally, together with the use and development of additional analyses, it will be necessary to further investigate the link between the presence of lexicon identifiers and psychological distress, possibly following specific users within existing communities through conversation graphs. This must be further investigated also considering the only partial overlap between the available findings of the top-down and bottom-up approaches.

CRediT authorship contribution statement

Marco Viviani: Conceptualization, Data curation, Formal analysis, Methodology, Resources, Supervision, Validation, Writing - original draft, Writing - review & editing. **Cristina Crocamo:** Conceptualization, Data curation, Formal analysis, Methodology, Investigation, Resources, Writing - review & editing. **Matteo Mazzola:** Methodology, Software, Data curation, Validation, Visualization. **Francesco Bartoli:** Conceptualization, Investigation, Resources, Writing - review & editing. **Giuseppe Carrà:** Conceptualization, Supervision, Writing - review & editing. **Gabriella Pasi:** Conceptualization, Supervision, Writing - review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data and code availability

Data, code, and documentation are available at the following URL: https://github.com/ikr3-lab/vulnerability_analysis/

Funding

This work received no financial support.

References

[1] K. Garimella, G.D.F. Morales, A. Gionis, M. Mathioudakis, Quantifying controversy on social media, ACM Trans. Soc. Comput. 1 (1) (2018) 1–27.

[2] M. Lai, M. Tambuscio, V. Patti, G. Ruffo, P. Rosso, Stance polarity in political debates: A diachronic perspective of network homophily and conversations on Twitter, Data Knowl. Eng. 124 (2019) 101738.

[3] R. Sicilia, S.L. Giudice, Y. Pei, M. Pechenizkiy, P. Soda, Twitter rumour detection in the health domain, Expert Syst. Appl. 110 (2018) 33–40.

[4] E.O. Nsoesie, L. Flor, J. Hawkins, A. Maharana, T. Skotnes, F. Marinho, J.S. Brownstein, Social media as a sentinel for disease surveillance: what does sociodemographic status have to do with it? PLoS Curr. 8 (2016).

[5] M. Conway, D. O'Connor, Social media, big data, and mental health: current advances and ethical implications, Curr. Opin. Psychol. 9 (2016) 77–82.

[6] M. Dredze, How social media will change public health, IEEE Intell. Syst. 27 (4) (2012) 81–84.

[7] C. Crocamo, M. Viviani, F. Bartoli, G. Carrà, G. Pasi, Detecting binge drinking and alcohol-related risky behaviours from Twitter's users: An exploratory content-and topology-based analysis, Int. J. Environ. Res. Public Health 17 (5) (2020) 1510.

[8] R.B. Correia, I.B. Wood, J. Bollen, L.M. Rocha, Mining social media data for biomedical signals and health-related behavior, Annu. Rev. Biomed. Data Sci. 3 (2020).

[9] J.C. Eichstaedt, H.A. Schwartz, M.L. Kern, G. Park, D.R. Labarthe, R.M. Merchant, S. Jha, M. Agrawal, L.A. Dziurzynski, M. Sap, et al., Psychological language on Twitter predicts county-level heart disease mortality, Psychol. Sci. 26 (2) (2015) 159–169.

[10] K. Kircaburun, Self-esteem, daily internet use and social media addiction as predictors of depression among turkish adolescents, J. Educ. Pract. 7 (24) (2016) 64–72.

[11] T. Moretta, G. Buodo, Modeling problematic Facebook use: Highlighting the role of mood regulation and preference for online social interaction, Addict. Behav. 87 (2018) 214–221.

[12] É. Verseillié, S. Laconi, H. Chabrol, Pathological traits associated to Facebook and Twitter among french users, Int. J. Environ. Res. Public Health 17 (7) (2020) 2242.

[13] M. De Choudhury, S. Counts, E. Horvitz, Social media as a measurement tool of depression in populations, in: Proceedings of the 5th Annual ACM Web Science Conference, 2013, pp. 47–56.

[14] D. Mowery, H. Smith, T. Cheney, G. Stoddard, G. Coppersmith, C. Bryan, M. Conway, Understanding depressive symptoms and psychosocial stressors on Twitter: a corpus-based study, J. Med. Internet Res. 19 (2) (2017) e48.

[15] Y. Hswen, J.A. Naslund, J.S. Brownstein, J.B. Hawkins, Online communication about depression and anxiety among twitter users with schizophrenia: preliminary findings to inform a digital phenotype using social media, Psychiatr. Q. 89 (3) (2018) 569–580.

[16] Y. Hswen, A. Gopaluni, J.S. Brownstein, J.B. Hawkins, Using twitter to detect psychological characteristics of self-identified persons with autism spectrum disorder: a feasibility study, JMIR MHealth UHealth 7 (2) (2019) e12264.

[17] C. Chew, G. Eysenbach, Pandemics in the age of Twitter: content analysis of Tweets during the 2009 H1N1 outbreak, PLoS One 5 (11) (2010) e14118.

[18] A.A. Anderson, Expressions of resilience: social media responses to a flooding event, Risk Anal. (2020).

[19] D. Garcia, B. Rimé, Collective emotions and social resilience in the digital traces after a terrorist attack, Psychol. Sci. 30 (4) (2019) 617–628.

[20] J. Xue, J. Chen, R. Hu, C. Chen, C. Zheng, Y. Su, T. Zhu, Twitter discussions and emotions about the COVID-19 pandemic: Machine learning approach, J. Med. Internet Res. 22 (11) (2020) e20550.

[21] X. Ji, S.A. Chun, Z. Wei, J. Geller, Twitter sentiment classification for measuring public health concerns, Soc. Netw. Anal. Min. 5 (1) (2015) 13.

[22] R. Chandrasekaran, V. Mehta, T. Valkunde, E. Moustakas, Topics, trends, and sentiments of tweets about the COVID-19 pandemic: Temporal infoveillance study, J. Med. Internet Res. 22 (10) (2020) e22624.

[23] C. Crocamo, M. Viviani, L. Famiglini, F. Bartoli, G. Pasi, G. Carrà, Surveilling COVID-19 emotional contagion on Twitter by sentiment analysis, Eur. Psychiatry 64 (1) (2021).

[24] N.A. Lukianova, Y.B. Burkatovskaya, E.V. Fell, Sten score method and cluster analysis: Identifying respondents vulnerable to drug abuse, in: The European Proceedings of Social & Behavioural Sciences (EpSBS). Vol. 35: Research Paradigms Transformation in Social Sciences, Vol. 352017, (RPTSS 2017).—Nicosia, 2018, Future Academy, 2018, pp. 779–789.

[25] B.L. Curtis, S.J. Lookatch, D.E. Ramo, J.R. McKay, R.S. Feinn, H.R. Kranzler, Meta-analysis of the association of alcohol-related social media use with alcohol consumption and alcohol-related problems in adolescents and young adults, Alcoholism Clin. Exp. Res. 42 (6) (2018) 978–986.

[26] T. Singh, K. Roberts, T. Cohen, N. Cobb, J. Wang, K. Fujimoto, S. Myneni, Social media as a research tool (SMaaRT) for risky behavior analytics: methodological review, JMIR Public Health Surv. 6 (4) (2020) e21660.

[27] R.M. Merchant, E.C. South, L. Nicole, Public health messaging in an era of social media, J. Am. Med. Assoc. (2021).

[28] X. Huang, Z. Li, Y. Jiang, X. Li, D. Porter, Twitter reveals human mobility dynamics during the COVID-19 pandemic, PLoS One 15 (11) (2020) e0241957.

[29] M.O. Lwin, J. Lu, A. Sheldenkar, P.J. Schulz, W. Shin, R. Gupta, Y. Yang, Global sentiments surrounding the COVID-19 pandemic on Twitter: analysis of Twitter trends, JMIR Public Health Surv. 6 (2) (2020) e19447.

[30] D. Valdez, M. Ten Thij, K. Bathina, L.A. Rutter, J. Bollen, Social media insights into US mental health during the COVID-19 pandemic: Longitudinal analysis of Twitter data, J. Med. Internet Res. 22 (12) (2020) e21418.

[31] S.C. Guntuku, G. Sherman, D.C. Stokes, A.K. Agarwal, E. Seltzer, R.M. Merchant, L.H. Ungar, Tracking mental health and symptom mentions on twitter during covid-19, J. Gen. Intern. Med. 35 (9) (2020) 2798–2800.

[32] J. Hussain, F.A. Satti, M. Afzal, W.A. Khan, H.S.M. Bilal, M.Z. Ansaar, H.F. Ahmad, T. Hur, J. Bang, J.-I. Kim, et al., Exploring the dominant features of social media for depression detection, J. Inf. Sci. 46 (6) (2020) 739–759.

[33] H.M. Wallach, Topic modeling: beyond bag-of-words, in: Proceedings of the 23rd International Conference on Machine Learning, 2006, pp. 977–984.

[34] D.M. Blei, Probabilistic topic models, Commun. ACM 55 (4) (2012) 77–84.

[35] B.V. Barde, A.M. Bainwad, An overview of topic modeling methods and tools, in: 2017 International Conference on Intelligent Computing and Control Systems, ICICCS, IEEE, 2017, pp. 745–750.

[36] H. Jelodar, Y. Wang, C. Yuan, X. Feng, X. Jiang, Y. Li, L. Zhao, Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey, Multimedia Tools Appl. 78 (11) (2019) 15169–15211.

[37] J. Chuang, C.D. Manning, J. Heer, Termite: Visualization techniques for assessing textual topic models, in: Proceedings of the International Working Conference on Advanced Visual Interfaces, 2012, pp. 74–77.

[38] C. Sievert, K. Shirley, LDAvis: A method for visualizing and interpreting topics, in: Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces, 2014, pp. 63–70.

[39] M. Röder, A. Both, A. Hinneburg, Exploring the space of topic coherence measures, in: Proceedings of the Eighth ACM International Conference on Web Search and Data Mining, 2015, pp. 399–408.

[40] Y.R. Tausczik, J.W. Pennebaker, The psychological meaning of words: LIWC and computerized text analysis methods, J. Lang. Soc. Psychol. 29 (1) (2010) 24–54.

[41] D.E. Losada, F. Crestani, A test collection for research on depression and language use, in: N. Fuhr, P. Quaresma, T. Gonçalves, B. Larsen, K. Balog, C. Macdonald, L. Cappellato, N. Ferro (Eds.), Experimental IR Meets Multilinguality, Multimodality, and Interaction, Springer International Publishing, Cham, 2016, pp. 28–39.

[42] C. Hutto, E. Gilbert, VADER: A parsimonious rule-based model for sentiment analysis of social media text, 2014, URL https://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/view/8109.

[43] M. Müller, M. Salathé, P.E. Kummervold, COVID-Twitter-BERT: A natural language processing model to analyse COVID-19 content on Twitter, 2020, arXiv preprint arXiv:2005.07503.

[44] B. Verma, R.S. Thakur, Sentiment analysis using lexicon and machine learning-based approaches: A survey, in: Proceedings of International Conference on Recent Advancement on Computer and Communication, Springer, 2018, pp. 441–447.

[45] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, Bert: Pre-training of deep bidirectional transformers for language understanding, 2018, arXiv preprint arXiv:1810.04805.

[46] S. Rosenthal, N. Farra, P. Nakov, SemEval-2017 task 4: Sentiment analysis in Twitter, in: Proceedings of the 11th International Workshop on Semantic Evaluation, SemEval-2017, 2017, pp. 502–518.

Marco Viviani is an Assistant Professor at the University of Milano-Bicocca (Milan, Italy), Department of Informatics, Systems, and Communication (DISCo). He works in the Information and Knowledge Representation, Retrieval and Reasoning (IKR3) Lab. His main research interests include Social Computing, Social Media Analytics, Information Retrieval, Text Mining, Trust and Reputation Management, User Modeling.
He has been Co-chair of several Special Tracks and Workshops at International Conferences, General Co-chair of MDAI 2019, and Guest Editor of several Special Issues in International Journals. He serves as a peer reviewer for more than 30 International Journals, and he is Program Committee Member of many International Conferences. He has published more than 100 research works as Internationl Conference papers, articles in International Journals, Book Chapters and Monographs.

Cristina Crocamo is a Postdoctoral Researcher at the University of Milano-Bicocca (Milan, Italy), working also as advisor for medical data management and analysis. Her research interests include Biostatistics, Digital Health, and Public Mental Health, along with Psychiatric Epidemiology, Addiction Psychiatry, and Statistical Modelling of the Pathophysiology of Mental Disorders. She has also expertise in systematic reviews and meta-analyses with extensive knowledge of statistical methods. She has collaborated in both European and national projects. She has co-authored several publications in Peer-reviewed Journals, and she is Section Board Member of the International Journal of Environmental Research and Public Health (Public Health Statistics and Risk Assessment Section).

Matteo Mazzola is currently working as a software developer at SpindoxLabs, a Research and Development Lab located in Trento (Italy). He holds a Master's Degree in Data Science from the University of Milano-Bicocca (Milan, Italy), where he collaborated on some research projects. His research interests include Natural Language Processing, Text Mining, and Graph Theory.
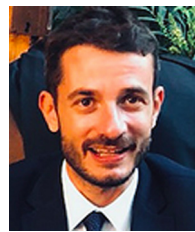
Francesco Bartoli is an Assistant Professor in Psychiatry at the University of Milano-Bicocca (Milan, Italy). He has expertise in meta-analyses and observational studies in the fields of Biological Psychiatry, Psychopharmacology, and Psychiatric Epidemiology. He is a Board Member of the Early Career Psychiatrists section of the Italian Society of Psychiatry. He is Associate Editor of 'BMC Psychiatry', Review Editor of 'Frontiers in Psychiatry', Editorial Board Member of 'Alpha Psychiatry', and Section Board Member of the 'International Journal of Environmental Research and Public Health'. He served as a peer reviewer for more than 50 International Journals. He won the 2020 'Brain Sciences' Young Investigator Award for his research in this field.

Giuseppe Carrà is an Associate Professor of Psychiatry at the University of Milano-Bicocca (Milan, Italy), and a Research Associate at the Division of Psychiatry, University College of London (UK).
Giuseppe extensively researched in the addiction field in the last 25 years, with particular focus on comorbid severe mental illnesses and substance abuse.
Author of over 180 papers in scientific, Peer-reviewed Journals, Giuseppe co-edited the 'Textbook of Addiction Treatment: International Perspectives', Springer-Nature 2015. Second edition, Springer-Nature 2020. Clinically, he works as the Medical Director of a large Mental Health and Addiction Department in Milan (I).

Gabriella Pasi is a Full Professor at the University of Milano-Bicocca (Milan, Italy), Department of Informatics, Systems, and Communication (DISCo). She is currently Head of Department and Pro-rector for International Relations at the University of Milano-Bicocca. Since 2005 she leads the Information and Knowledge Representation, Retrieval and Reasoning (IKR3 Lab), previously Information Retrieval Laboratory (IR Lab), at DISCo. Her main research interests are related to Information Retrieval, Information Filtering and Decision Support. In recent years she has addressed the issues of Contextual Search and User Modelling. She is also conducting research activities related to the analysis of User-Generated Content on social media. She has been Program chair and General Co-chair of several International Conferences, she is Associate Editor of International Journals (e.g., ACM Computing Surveys and IEEE Transactions of Fuzzy Systems) and she has published more than 300 papers in prestigious venues (both International Journals and proceedings of International Conferences). From 2013 to 2017 she has been president of the European Association on Fuzzy Logic and Technologies (EUSFLAT).