JMIR Publications
The Leading eHealth Publisher

Social media concerns during the COVID-19 Pandemic: An analysis of Sina Weibo Microblogging data

Junze Wang, Ying Zhou, Wei Zhang, Richard Evans, Chengyan Zhu

Submitted to: Journal of Medical Internet Research
on: July 05, 2020

Ahead of Print
JMIR Publications

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on it's website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressively prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Ahead of Print
JMIR Publications

Social media concerns during the COVID-19 Pandemic: An analysis of Sina Weibo Microblogging data

Junze Wang[1, 2] PhD; Ying Zhou[1, 2] BA; Wei Zhang[3] PhD; Richard Evans[4] PhD; Chengyan Zhu[5] PhD

[1]Non-traditional Security Center Huazhong University of Science and Technology Wuhan CN
[2]College of Public Administration Huazhong University of Science and Technology Wuhan CN
[3]School of Medicine and Health Management Huazhong University of Science and Technology Wuhan CN
[4]College of Engineering, Design and Physical Sciences Brunel University London London GB
[5]School of Political Science and Public Administration Wuhan University Wuhan CN

Corresponding Author:
Wei Zhang PhD
School of Medicine and Health Management
Huazhong University of Science and Technology
13# Hangkong Road
Qiaokou District
Wuhan
CN

Proof of Print
JMIR Publications

Abstract

Background: The COVID-19 pandemic has created a global health crisis, affecting economies and societies around the world. During times of uncertainty and unexpected change, citizens have turned to social media as a mediating tool for communicating primary sources of information. Platforms have allowed communities to come together for discussion and emotional support, and have played an important role for individuals, governments and organizations in publishing and exchanging information, and expressing opinions. However, limited research exists that studies the main concerns of citizens on social media during the pandemic.

Objective: The purpose of this study is to examine the main concerns raised and discussed by citizens on Sina Weibo, the largest social media platform in China, during the COVID-19 pandemic.

Methods: This study uses a web crawler tool and a set of predefined search terms ("New Crown Pneumonia," "New Crown Virus" and "COVID-19") to investigate the concerns raised on the Sina Weibo platform. We collect the textual information and metadata (number of likes, comments, retweets, publishing time, and publishing location) of microblogs published from February 2, 2020 to March 15, 2020. After segmenting the words of the text collected, we used a topic modeling technique, the Latent Dirichlet Allocation (LDA) algorithm, to identify the most common topics posted by microbloggers. We analyzed the emotional tendency of topics, calculated the proportion distribution of topics, performed user' behavior analysis on the topics using the data of: number of likes, comments and retweets, and studied the change in user concerns and the differences in participatory discussion between citizens living in different regions in China.

Results: Based on the 158,386 eligible microblog posts collected, we identified 11 topics, which were grouped into 6 themes. Ranked by the level of attention, these topics were: Pandemic statistics (27.7%); Transmission routes and preventive measures (22.7%); Epidemics in Wuhan (16.5%); Detection and quarantine (13.4%); New Crown Treatments (13.1%); Work and production resumption (12.9%); Epidemics in other countries around the world (11.7%); refueling and saluting towards anti-epidemic action (10.1%); Economic impact (6%); and Study resumption (3.5%). The mean sentiment was positive for 6 topics and negative for 5 topics. The topic with the highest mean of retweets was: Epidemics in Wuhan, while the topic with the highest mean of comments and likes was: Study resumption.

Conclusions: Citizen's concerns on social media are highly correlated with the development of the pandemic. During the COVID-19 pandemic, social media has provided a platform for Chinese government departments and organizations to better understand public concerns and demands. Similarly, social media has provided a channel to disseminate information about epidemic prevention and has influenced public' attitudes and behaviors. Government departments, especially health related, can create appropriate policies in a timely manner through monitoring social media platforms to guide the public's views and behaviors during epidemics.

(JMIR Preprints 05/07/2020:22152)
DOI: https://doi.org/10.2196/preprints.22152

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

Please make my preprint PDF available to anyone at any time (recommended).

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.

✓ **Only make the preprint title and abstract visible.**

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

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

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

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

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

Ahead of Print
JMIR Publications

Original Manuscript

Ahead of Print
JMIR Publications

Social media concerns during the COVID-19 Pandemic: An analysis of Sina Weibo Microblogging data

Abstract:

Background: The COVID-19 pandemic has created a global health crisis, affecting economies and societies around the world. During times of uncertainty and unexpected change, citizens have turned to social media as a mediating tool for communicating primary sources of information. Platforms, such as Twitter and Sina Weibo, have allowed communities to come together for discussion and emotional support, while playing an important role for individuals, governments and organizations in publishing and exchanging information and expressing opinions. However, limited research exists which studies the main concerns of citizens on social media during the pandemic.

Objective: To examine the main concerns raised and discussed by citizens on Sina Weibo, the largest social media platform in China, during the COVID-19 pandemic.

Method: This study uses a web crawler tool and a set of predefined search terms (“New Crown Pneumonia”, “New Crown Virus” and “COVID-19”) to investigate concerns raised by users of the Sina Weibo platform. Textual information and the metadata (number of likes, comments, retweets, publishing time, and publishing location) of microblog posts, published between 1 December 2019 to 31 July 2020, were collected. After segmenting the words of the text collected, we used a topic modeling technique, Latent Dirichlet Allocation (LDA), to identify the most common topics posted by users. We analyzed the emotional tendency of topics, calculated the proportional distribution of topics, performed user’ behavior analysis on the topics using data collected from the number of likes, comments and retweets, and studied the change in user concerns and differences in participation between citizens living in different regions of Mainland China.

Results: Based on the 203,191 eligible microblog posts collected, we identified 17 topics which were grouped into 8 themes. Ranked by the level of attention paid to each, these topics were: Pandemic statistics (26.8%); Domestic epidemic (18.1%); Epidemics in other countries around the world (13.9%); New Crown treatments (11.1%); Medical resources (10.9%); Economic shock (10.1%); Quarantine and investigation (9.12%); Patients’ cry for help (8.61%); Work and production resumption (7.83%); Psychological influence (7.12%); Joint Prevention and Control (6.13%); Material donation (6.11%); Epidemics in neighboring countries (5.90%); Vaccine development (4.64%); Refueling and saluting towards anti-epidemic action (4.63%); Detection (4.39%); and Study resumption (3.39%). The mean sentiment was positive for 11 topics and negative for 6 topics. The topic with the highest mean of retweets was ‘Domestic epidemic’, while the topic with the highest mean of likes was ‘Quarantine and investigation’.

Ahead of Print © JMIR Publications

Conclusions: Citizens' concerns on social media are highly correlated with the evolution of the global pandemic. During the COVID-19 pandemic, social media has provided a platform for Chinese government departments and organizations to better understand public concerns and demands. Similarly, social media has provided a channel to disseminate information about epidemic prevention and has influenced public attitudes and behaviors. Government departments, especially those which are health related, can create appropriate policies in a timely manner through the monitoring of social media platforms to guide public opinions and behaviors during epidemics.

Keywords: Coronavirus; COVID-19; Social Media; Public Health; Sina Weibo; Public opinion; Citizen concerns.

1.Introduction

On 29 June 2020, The World Health Organization (WHO) marked the six-month anniversary of the COVID-19 outbreak. The first case of the unknown pneumonia strain, now known as COVID-19 or SARS-CoV-2, was reported in Wuhan City, Hubei Province, on 31 December 2019 and has subsequently made a profound impact around the world. After clusters of pneumonia cases were reported in Wuhan in early January 2020, laboratory analyses were conducted which led to the epidemic being identified as a new coronavirus, officially named COVID-19 by the WHO. By the 3rd of July 2020, 83,545 confirmed cases were reported in Mainland China [1]. After the outbreak of COVID-19 in China, it quickly spread globally. At the end of 18th August 2020, there were a total of 214 countries and regions with confirmed cases of COVID-19 worldwide, with the total number of contracted people exceeding 21 million and the total number of deaths worldwide exceeding 770,000. The global average mortality rate of COVID-19 is about 0.3%-1.5%, but the mortality rate of COVID-19 in countries, such as the United States, Brazil and Mexico, is much higher than this. With the spread of COVID-19 around the world, the epidemic gradually attracted widespread attention and discussion on social media platforms. This trend was particularly evident on Sina Weibo, one of the largest social media platforms in China.

Extant studies have demonstrated that timely understanding of public attitudes and demands plays an important role in responding to public crises [2,3]. Most social media platforms possess media-oriented features that are crucial to mediating information dissemination. Citizens can receive the latest, fact-checked information provided by governments on social media in a timely manner, while governments can use the publicly available information released by citizens to better understand public attitudes, concerns and demands [4–6]. Sina Weibo, with over 500 million users by May 2020, provides a variety of communication mechanisms for citizen interaction, allowing the

Ahead of Print
JMIR Publications

Chinese public to share information and exchange opinions [7–9]. By analyzing the concerns raised by citizens about COVID-19 in microblog posts shared on Sina Weibo, governments can better understand public attitudes and demands [10,11], and clarify existing challenges in government departments and organizations when dealing with the pandemic. This research provides important insights and implications for policy makers, especially those working in public health departments. The results offer deeper understanding of public perception and highlight shortcomings in practice to better meet public needs.

A large body of literature exists that has examined the role of social media for analyzing public behavior, attitudes and responses during times of public crises. Researchers have mainly focused on platform functionality, user behavior characteristics, and the use of social media during crises. As most platforms are freely available in the public domain [12], they have become the widely adopted method for citizens to stay connected, discuss concerns and opinions, and escape the monotony of lockdown during the COVID-19 pandemic [13]. Researchers agree that social media has become an important medium for information dissemination during the pandemic [14] and plays a unique role in information sharing [15] and healthcare discussion [16]. Some scholars have explored user behavior and connection networks on social media platforms and believe that personality characteristics affect users' behavior [17], while networks comprise both positive and negative connections [18,19]. The unique position and role of social media in response to public crises have also attracted major attention among researchers. They concur that platforms play an important role during crisis management, especially in terms of providing citizens with timely information [20] and reducing citizens' anxiety and fear [21]. Based on existing research, this paper collects data from the leading Chinese microblogging platform, Sina Weibo, to analyze the main concerns expressed by citizens during the COVID-19 pandemic.

2. Methodology

2.1 Data Collection

Sina Weibo is the leading microblogging platform in China which allows users to send and receive short character-limited posts and retrieve textual content by searching for specified keywords over a defined date range. Using this functionality, we collected shared microblog posts, relating to COVID-19, during the time period of 1 December 2019 to 31 July 2020. The web crawler tool, Octopus, was used to search for predefined keywords, including: "New Crown Pneumonia", "Coronavirus", "New Crown Virus" and "COVID-19". In addition to the textual content collected, we sourced the metadata for each microblogged post, including number of likes, number of comments, number of retweets, publishing time, and publishing location. The location information in

this data refers to the registered address of the Sina Weibo account that posted the microblog. In order to obtain this data, we used the advanced search functionality of Sina Weibo.

2.2 Data Preprocessing

In the Chinese language, there is no obvious separation between words. Therefore, for the purpose of completing the LDA processing tasks, it was necessary to add obvious separators between words. We performed Chinese segmentation on the textual content collected during data preprocessing. The widely used Chinese word segmentation tool, ictclas, was employed to divide microblog posts into groups of words separated by spaces. For example, the microblog post “JD announced the donation of 1 million medical masks and 60,000 pieces of medical supplies to Wuhan City in batches” was segmented into “JD / announced / the / donation of / 1 million medical masks / and / 60,000 pieces of / medical supplies / to / Wuhan City / in batches”. In addition, we removed stop words [22] in texts, which is similar to removing stop words, such as ‘an’ and ‘the’, in English text analysis. Chinese stop words mainly include the following two categories: (1) widely used vocabulary, such as: me, you, some and every day, and (2) auxiliary words for mood, adverbs, prepositions, conjunctions and other words that have no meaning by themselves, such as: in, yes, so, and then [23].In addition, punctuation and characters, such as emojis, were removed.

2.3 Data Analysis

We applied topic modeling by specifying the number of topics required by the LDA to separate the set of microblog posts into defined clusters[24]. The LDA technique can be used to identify the most common topics in microblog posts shared on Sina Weibo. Topic modeling is an unsupervised machine learning technique that can identify clusters in a collection of documents (microblog posts in our case). In this study, we used the latent Dirichlet allocation algorithm from the LDA4j package. LDA4j is implemented in the JAVA language of the LDA algorithm; the project can be downloaded for free from github: https://github.com/hankcs/LDA4j.

LDA is a widely used topic modeling algorithm [25]. According to the LDA model, a document (i.e., the text of a microblog post) is a collection of vocabularies and may include multiple topics. The goal of the LDA model is to speculate the distribution of topics, based on a given document [26]. With LDA modeling, we can map the given documents to a fixed set of topics and capture the representative words for each topic. Then, the natural clusters in the microblog posts dataset can be established.

In order to determine the appropriate number of LDA topics, we used this coherence score to draw judgments, which is a method proposed for selecting a suitable number of LDA topics[27]. Through continuous adjustment of the number of topics, we found that when the number of topics

Ahead of Print JMIR Publications

was 17, the consistency score was at its optimal value[28]. We, therefore, set the LDA model to separate the set of microblog posts into 17 clusters; see text sets in Multimedia Appendix 1.

Subsequently, we conducted manual analysis and selected representative and high-proportion keywords from the top 30 keywords of each topic. Next, consensus was reached on the 17 topics and related keywords; the topics and their keywords can be seen in Table 2. Lastly, we used these keywords to classify the microblog posts; we also obtained the number of microblog posts under each topic and the proportion of each topic in all related microblogs; see examples of microblog posts for each topic in Multimedia Appendix 2.

For example, “The sudden outbreak of new coronal pneumonia has had an impact on the country’s economic operation. It has brought a negative GDP growth and an increasingly complicated international and domestic environment. However, under the strong hedging of the counter-cyclical adjustment policy, the resumption of production has advanced rapidly, while the main economic indicators show a rebound in March, and the decline rate narrowed significantly.” This microblog can be categorized into two different topics: “Work and production resumption” and “Economic shock”.

We also performed other analyses of the collected data, such as sentiment analysis. The sentiment score varied between –1.0 to 1.0, with –1.0 being the most negative text and 1.0 being the most positive text. In addition, we calculated the interaction rate of users for each topic by analyzing the average number of retweets, likes and comments per topic. Finally, using the publishing time and location, we analyzed changes in user concerns, based on time periods, and differences in user engagement in discussions related to COVID-19, based on Chinese regions.

3. Results

Using the web crawler tool and predefined search terms, we obtained a total of 203,191 microblog posts from the Sina Weibo platform, which were shared between 1 December 2019 and 31 July 2020.

3.1 Microblog Analysis

3.1.1 Topics

According to the results obtained from the LDA and the keywords involved in each topic, we were able to group topics into 8 Themes, including: (1) Patient admission; (2) Treatment and research; (3)Treatment resources; (4) Fighting the epidemic together; (5) Order restoration work; (6) Prevention and control measures of COVID-19; (7) Domestic and overseas pandemic situation; and (8) Impact of COVID-19. The corresponding terms for each topic are shown in Table 1; these

keywords are also used as the criteria for topic classification.

Table 1. Topics covered in microblog posts and the representative terms corresponding to each topic

Theme	Topics	Terms corresponding to each topic
Patient admission	Patients' cry for help	help, attention, diffusion, receiving
Treatment and research	New Crown treatments	discharged from the hospital, cure, treatment, rehabilitation, Chinese Medicine
	Vaccine development	vaccines, prevention, drugs, clinical trials.
Treatment resources	Medical resources	medical treatment, doctor, medical care, nurse, mobile cabin, ward.
Fighting the epidemic together	Material donation	mask, materials, donations
	Refueling and saluting towards anti-epidemic action	work with one heart, saluting, unity is strength, responsibility, persevere, fight
Order restoration work	Work and production resumption	resumption of work, resumption of production, employment, operation
	Study resumption	return to school, student, school, high school, college entrance examination
Prevention and control measures of COVID-19	Quarantine and investigation	quarantine, contact, fever, 14 days, investigation
	Joint prevention and control	headquarters, work leading group, joint defense, joint control
	Detection	nucleic acid, positive, negative
Domestic and overseas pandemic situation	Domestic epidemic	Wuhan, Hubei, prison, Zhejiang, Shandong
	Epidemics in neighboring countries	Japan, South Korea, Tokyo, Russia
	Epidemics in other countries around the world	America, Trump, Britain, Italy, India, Brazil, France
	Epidemic statistics	confirmed cases, new cases, cumulative cases, suspected cases
Impact of COVID-19	Economic shock	economy, influence, market, shock
	Psychological influence	hope, worry, fear, terrible

3.1.2 Theme 1: Patient admission

The topic contained in this theme is Patients' cry for help, which relates to patients that have or might have contracted COVID-19 and their treatment situation. On the one hand, the cry for help of these patients who are awaiting treatment has attracted widespread attention. On the other hand, the subsequent reception and treatment of these patients has also attracted significant attention.

3.1.3 Theme 2: Treatment and research

This theme contains two topics. The first topic is New Crown treatment. Under this topic, the patients' cure has received widespread attention, such as the number of people discharged from hospital and the number of people cured. At the same time, treatment methods have also attracted widespread public attention, such as treatment using Chinese medicine. The second topic is Vaccine development. In this topic, research progress in relation to vaccines and their clinical trials have

attracted much comment. Meanwhile, there has been widespread debate about whether vaccines can achieve the goal of preventing the virus.

3.1.4 Theme 3: Treatment resources

The topic under this theme is medical resources. This topic focuses on core aspects of medical resources; medical staff and hospital wards, the allocation and integration of medical resources, and the establishment of temporary hospitals, such as mobile cabin hospitals – these have all been of general concern.

3.1.5 Theme 4: Fighting the epidemic together

This theme involves two topics. The first one is Material donation. In the early stages of fighting the epidemic in Mainland China, masks and other anti-epidemic materials were extremely scarce. This issue aroused widespread concern and the donation of various anti-epidemic materials has become a hot topic. The second topic is Refueling and saluting towards anti-epidemic action. This includes two aspects: first, the confidence and determination to beat the pandemic. Keywords, such as “fighting” and “defeat”, represent the general attitude of Sina Weibo users towards the fight against COVID-19. Secondly, cooperation in the fight against the pandemic is often discussed; keywords, such as “work with one heart” and “unity is strength”, have frequently been mentioned.

3.1.6 Theme 5: Order restoration work

As the spread of COVID-19 in Mainland China is gradually controlled, how to restore production and resume normal ways of living has become a hot topic, which includes two sub-topics. The first relates to work and production resumption. The progress and arrangements for the resumption of work and production have aroused widespread concern and discussion. This work, which is aimed at restoring normal production and operational order, has received widespread support from the Chinese public. The second topic is study resumption, which mainly involves two aspects. First, the time when students of all ages will return to school. Second, the time for the national college entrance examination.

3.1.7 Theme 6: Prevention and control measures of COVID-19

There are three topics included in this theme. The first one is Quarantine and investigation, which includes two aspects. First, the investigation of symptoms, such as coughs and fever. Second, quarantine periods – in particular, the 14-day quarantine period has become standard.

The second one is Joint prevention and control. Taking the measures of joint prevention and control and establishing a headquarters is an important action taken by the Chinese government in response to the epidemic. This action is also key for the Chinese government to contain the epidemic in the short term. Correspondingly, it has also been widely recognized and been of great concern to

Ahead of Print © JMIR Publications

the public.

The last topic is Detection. The main content under this topic is nucleic acid detection which is an important way to establish who are or who have been infected by COVID-19. The topic of detection has aroused widespread discussion.

3.1.8 Theme 7: Domestic and overseas pandemic situation

Four topics are included under this theme. The first is Domestic epidemic. On the one hand, as the main battlefield in the fight against the epidemic in China, the epidemic situation in Hubei and Wuhan has received significantly high attention. On the other hand, the spread of the epidemic in China is also of great concern to the public. Epidemics in neighboring countries is the second topic. The peak of this topic appeared in mid-to-late February 2020, when the epidemic began to spread from China to neighboring countries, such as Japan, South Korea and Russia – this has aroused heated discussion among Sina Weibo users.

The third topic is Epidemics in other countries around the world. At the time when this topic appeared, the epidemic had spread across the globe, so the topic involved many countries and regions. The United States, as the area hit hardest by the epidemic, has attracted widespread attention from Sina Weibo microbloggers. In particular, Donald Trump's statement that "if (New Coronary Pneumonia) deaths can be controlled to less than 100,000, (this shows that) all of us are doing well" has been a matter of heated discussion. At the same time, other countries with serious numbers of cases, such as Italy, India, Brazil and France have also caused extensive debates. The fourth topic is Epidemic statistics. The epidemic data for COVID-19, such as the number of confirmed cases, new cases and suspected cases, received much attention and discussion, with the public's attention and discussion of the epidemic data running throughout all stages of the evolution of the epidemic. The epidemic statistics can be divided into two parts: domestic epidemic data and international epidemic data. Domestic epidemic data refers to case data released by the National Health Commission. For example, by the end of July 31, 2020, a total of 78,989 cases had been cured in China, and a total of 84,337 confirmed cases had been reported. International epidemic data is published by governments around the world. For example, by the end of July 31, 2020, the number of daily diagnosed cases globally exceeded 291,000, totaling 17.72 million. The United States alone added more than 70,000 diagnosed cases in one day.

3.1.9 Theme 8: Impact of COVID-19

This theme has two topics. First, Economic shock. This topic mainly refers to the negative impact of the epidemic on economies and markets around the world. Second, Psychological influence. On the one hand, the outbreak of the epidemic has delivered a huge psychological blow to

Ahead of Print
JMIR Publications

the public and emotions, such as worry and fear, have spread among citizens. On the other hand, the public also hopes and believes that the fight against the epidemic will be successful and they remain in optimistic mood.

3.2 Results of the Proportional Analysis of the Main Topics

To understand the specific degree of concern for the various topics identified, we established the number and proportion of each topic by counting the number of microblog posts per topic. For example, the proportion of each topic = the number of microblog posts shared about the topic / the total number of microblogs, 203,191. The results of our analysis are shown in Table 2.

Table 2 Number and proportion of microblog posts related to each topic

Sequence	Topic	Number of microblog posts	Proportion of microblog posts
1	Pandemic statistics	54513	26.8%
2	Domestic epidemic	36936	18.1%
3	Epidemics in other countries around the world	28329	13.9%
4	New Crown treatments	22725	11.1%
5	Medical resources	22246	10.9%
6	Economic shock	20699	10.1%
7	Quarantine and investigation	18541	9.12%
8	Patients' cry for help	17513	8.61%
9	Work and production resumption	15914	7.83%
10	Psychological influence	14485	7.12%
11	Joint prevention and control	12457	6.13%
12	Material donation	12434	6.11%
13	Epidemics in neighboring countries	11995	5.90%
14	Vaccine development	9428	4.64%
15	Refueling and saluting towards anti-epidemic action	9413	4.63%
16	Detection	8939	4.39%
17	Study resumption	6889	3.39%

As shown in Table 2, statistics relating to the pandemic are viewed as a metric to ascertain citizen understanding of the current situation during the epidemic. The topic has received significant attention since the initial outbreak and is ranked first, with 26.8%. The domestic epidemic situation has always been the focus of public attention, especially when the epidemic in Mainland China has not been fully controlled. The number of microblog posts on this topic accounted for 18.1%, ranking it second. As COVID-19 has spread around the world, microblog posts referring to the pandemic in other countries has grown consistently, with a percentage of 13.9%, ranked third.

The treatment of new coronary pneumonia has raised high expectations and received attention from the public for a long period with a percentage score of 11.1%. Medical resources are particularly important in the fight against the epidemic, which accounts for 10.9%. The huge impact of new crown pneumonia on the economy is closely related to the interests of the public, so the topic of economic shock also receives a high degree of attention; it is ranked sixth with 10.1%. As one of

Ahead of Print © JMIR Publications

the most important measures for controlling and preventing the spread of COVID-19, the topic of detection and quarantine has received considerable attention at all stages of the pandemic since its outbreak, accounting for 9.12% of total posts. The patient's cry for help has also attracted widespread attention, accounting for 8.61% of total posts.

An important measure to restore economic order in China is work and production resumption, which is closely related to citizens' lives. The percentage of microblog posts on this topic is 7.83%. The epidemic also greatly affected the public's psychology, so the topic of psychological impact has also received much attention, with a percentage of 7.12%. Joint prevention and control and material donations have received a certain amount of attention. These two topics ranked eleventh and twelfth, with 6.13% and 6.11%, respectively. The situation in neighboring countries has also aroused comment from Sina Weibo users, although the duration is short lived; accordingly, it ranks thirteenth with 5.90%.

The development of vaccines has also been discussed by the public, with related microblog posts accounting for 4.64%. The fight against COVID-19 is a national battle in which citizens refuel discussions on how to overcome the pandemic. The percentage of posts relating to this topic was 4.63%. As an important means of identifying patients that are infected with new coronary pneumonia, detection has also received a certain degree of attention, accounting for 4.39%. After the pandemic in Mainland China was controlled, study resumption was also mentioned, but compared to other topics, the degree of concern in this area is relatively low, accounting for only 3.39% of total posts.

3.3 Results of Users' Interaction and Sentiment Analysis

Based on the collected microblog metadata, including number of retweets, number of comments and number of likes, we were able to analyze the levels of interaction between users for each topic. At the same time, we performed a sentiment analysis for each topic to obtain the sentiment value for each. The values in Table 3 are the average number of retweets, comments, likes, and the sentiment values of each topic. The calculation formula to determine the sentiment value is (a-b)/(a-b), where a is the number of tweets expressing positive emotions, and b is the number of tweets expressing negative emotions [29].

Table 3 Results of interaction between users and sentiment analysis for each topic

Serial number	Topic	Retweets (mean)	Comments (mean)	Likes (mean)	Sentiment value
1	Work and production resumption	8	12	87	0.16
2	Epidemics in neighboring countries	29	22	199	-0.006
3	New Crown treatments	26	26	244	-0.05
4	Vaccine development	24	16	199	0.08
5	Pandemic statistics	15	12	178	-0.11
6	Economic shock	18	10	97	0.051
7	Material donation	35	17	245	0.058
8	Domestic epidemic	71	20	184	0.024
9	Medical resources	34	35	241	0.031
10	Quarantine and investigation	33	24	280	-0.049
11	Psychological influence	29	29	217	0.041
12	Joint Prevention and Control	15	13	236	0.146
13	Study resumption	7	10	129	0.083
14	Refueling and saluting towards anti-epidemic action	19	13	248	0.073
15	Epidemics in other countries around the world	11	10	118	-0.117
16	Patients' cry for help	22	18	224	-0.017
17	Detection	14	16	129	0.041

(Data description: The data reported on in this article was collected from microblogs posted by some opinion leaders. The number of retweets, comments and likes in such microblog posts are high, such as the microblog: "Zhong Nanshan said that Lianhuaqingwen capsules were proved to be effective in the treatment of COVID-19". The numbers of retweets, comments and likes for this microblog were: 13599, 19022 and 344780, respectively. Therefore, the average number of retweets, comments, and likes in this article is high.)

From Table 3, it may be seen that domestic epidemic, material donation and medical resources ranked as the top three in terms of the number of retweets, demonstrating that Sina Weibo users have a strong desire to share such information. By analyzing the number of likes, the mean is relatively high for the following topics: quarantine and investigation, refueling and saluting towards anti-epidemic action, material donation, and New Crown treatments. According to the sentiment value for each topic, the number of topics with positive sentiment values and negative sentiment values were 11 and 6 respectively. Furthermore, for most topics, the sentiment value was low, which means that the number of microblogs with positive emotions is almost the same as the number of microblogs with negative emotions. However, negative emotions are obvious, except for those related to pandemic statistics.

3.4 Results of Spatio-temporal Analysis on Microblogs

By using the publishing time of microblog posts, we were able to analyze the change in users'

concerns during different time periods. Similarly, we used the publishing location to analyze the differences in user engagement in discussions relating to COVID-19 between different regions in China. Figure 1 illustrates how the proportion of topics changed over the time period. The figure includes the names and proportions of the top four topics with the highest levels of attention from December 2019 to July 2020. Considering the different months, the top four topics constantly changed. From December 2019 to July 2020, there were 9 topics in the top four, including pandemic statistics, domestic epidemic, epidemics in other countries around the world, New Crown treatments, medical resources, economic shock, work and production resumption, psychological influence, and material donation. In December 2019 and January 2020, four topics received a high level of attention: domestic epidemic, pandemic statistics, psychological influence, and material donation. In February and March 2020, the domestic epidemic and pandemic statistics received more attention than before, with the number of microblog posts on these topics ranked first and second respectively. Meanwhile, the degree of attention paid to the topics “medical resources” and “New Crown treatments” increased, becoming part of the top four. In April and May 2020, resumption of work and production, and economic shock became hot topics in this time period. At the same time, pandemic statistics and epidemics in other countries around the world retained a high degree of attention. In June and July 2020, hot topics were little changed compared with previous months. Topics, such as pandemic statistics, epidemics in other countries around the world and economic shock, also maintained a high degree of attention during this period.

Figure 2 represents a geographical ranking of the number of microblogs. The location information is the registration address of the Weibo accounts which post the blogs. The total number of microblogs collected was 203,191, with the total number containing publishing location being 127,009, accounting for 62.5%. By analyzing the publishing location, we identified the top six regions with the largest number of microblog posts. It can be seen from Figure 2 that the Hubei Province received most posts, with its number being much higher than other provinces. Henan Province and Guangdong Province occupied second and third places with other provinces ranking below these.

Ahead of Print
©JMIR Publications

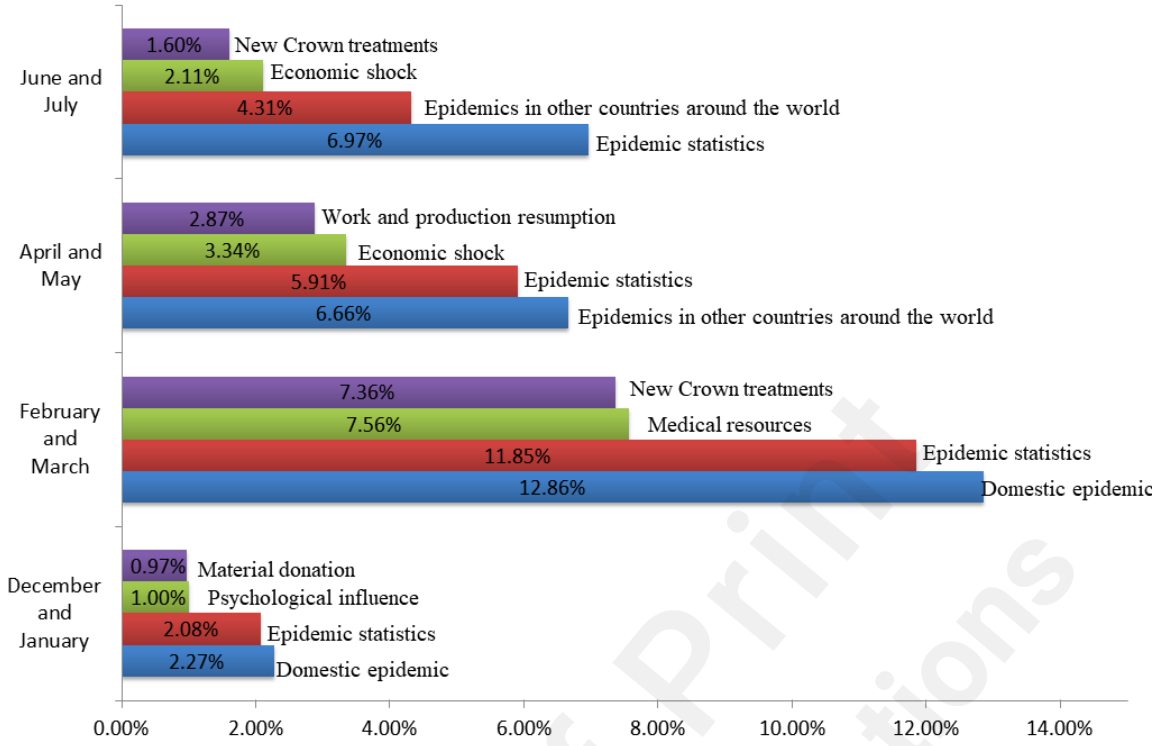

Figure 1. Changes in the proportion posts about the main topics from January to May

Ahead of Print
JMIR Publications

Figure 2. Number of microblogs by Chinese province

4. Discussion

4.1 Main Findings

Sina Weibo Users mainly focus on the latest news about new crown treatment and the status of the pandemic domestically and globally. Our study, conducted from 1 December 2019 to 31 July 2020, identified 8 themes with 17 topics. Among these topics, content related to COVID-19 treatment aroused high attention. In particular, content related to patients' outcry for help, and their emotions such as anxiety and panic were manifested. The widespread negative emotions on Sina Weibo make it easier for misinformation and rumors to influence public opinions, which negatively impacts anti-epidemic actions. In addition, Sina Weibo users have shown much interest in the progress of the treatment of new coronary pneumonia and the progress of vaccine development. At the same time, there has been considerable focus and discussion among a large number of Sina Weibo users on whether medical resources are sufficient and whether they can meet the needs of patients' treatment.

Sina Weibo users also focused heavily on the real-time situation of the epidemic at home and abroad. In particular, many microblog posts contained real-time data on the development of the

epidemic. Further, Sina Weibo users have not only paid attention to the epidemic situation in Chinese mainland, but have also discussed greatly the development and evolution of the epidemic in neighboring countries and the rest of the world. In addition, epidemic prevention and control measures, order restoration measures in the later stage of the epidemic, and the impact of the epidemic are also included in the 17 topics that Sina Weibo users have given attention.

The change in the number of microblog posts for the various topics in different time periods reflects the process of occurrence, development and demise of the pandemic in China. Our findings show that December 2019 and January 2020 were the key periods for responding to the pandemic, with domestic epidemic and pandemic statistics receiving high attention during this period. At the same time, the sudden outbreak also affected the public's psychology. On the one hand, the public believes that the fight against the epidemic will eventually be won; on the other hand, emotions such as fear and worry are also spreading among the public.

February and March 2020 were critical periods for curbing the spread of the epidemic in Mainland China [30]. During this period, pandemic statistics and the domestic epidemic still registered a high degree of concern. At the same time, the surge in the number of patients caused the public to pay attention to New Crown treatments and medical resources. From April to May 2020, the pandemic in Mainland China was basically under control, but it had begun to spread to other countries and regions. At this time, economic shock and work and production resumption became the most highly debated topics on Sina Weibo. Similarly, users paid great attention to the global spread of the pandemic, viewing real-time data.

From June to July 2020, the number of infections and deaths caused by COVID-19 worldwide continued to rise. The focus of Sina Weibo users shifted from China to the world and they began to pay greater attention to epidemics in other countries and the growing pandemic statistics.

The location of Sina Weibo users and whether they have experienced similar events are linked closely to their degree of attention to public crises. According to the spatial distribution information collected on posted microblogs, discussions related to COVID-19 differed among the various regions of China. The top 6 regions in the degree of participation were Hubei Province, Henan Province, Guangdong Province, Anhui Province, Jiangsu Province and Hunan Province. Hubei Province is the epicenter of COVID-19 in China, while Henan Province, Hunan Province and Anhui Province are all adjacent to Hubei Province, with many citizens working in Hubei Province. Jiangsu Province, as an economically developed region, has a relatively high mobility rate of personnel. As a result, Sina Weibo users in these provinces are more likely to be concerned about the pandemic. Guangdong Province experienced SARS in 2003, which is an important reason why the

Ahead of Print
JMIR Publications

region had more users who participated in discussions on the COVID-19 pandemic.

Sina Weibo and Twitter, both microblogging platforms, have similarities and differences between user concerns. A recent study reported the main topics that concern Twitter users about COVID-19 [29]**.** The authors analyzed four themes, containing 12 topics related to COVID-19 that Twitter users were concerned about from 2 February to 15 March 2020. They found that Twitter users were mainly concerned about the impact of coronavirus on people and countries. For example, the number of deaths related to coronavirus and the impact on citizens' emotions and psychology were mentioned in many tweets. In addition, the economic impact of COVID-19 has also been widely discussed. In particular, Twitter mentioned two main methods for reducing the spread of COVID-19: masks and quarantine. Compared with these findings, Sina Weibo users shared some similarities. Topics including economic impact and psychological influence have received high attention on both Sina Weibo and Twitter [29]. However, at the same time, the main concern of Sina Weibo users also demonstrate their own characteristics. First, Twitter users focus on the causes and effects of the epidemic, while Sina Weibo users pay greater attention to prevention, control and treatment. Various epidemic prevention and control measures, such as isolation, detection, and joint prevention and control, have attracted widespread attention among Sina Weibo users. Similarly, Sina Weibo users are highly concerned about content related to the new crown treatments, such as patients' outcry for help, medical resources, and treatment methods. In addition, citizens have continued to pay a high degree of attention to the development of the novel coronavirus vaccine. This may be attributed to the fact that the Chinese government has adopted a series of response measures, making citizens aware of the severity and harmfulness of the pandemic, meaning that they are more concerned about the health-related topics, such as the prevention, control and treatment of the virus.

Second, Sina Weibo users focus on the real-time status of the epidemic. This includes not only the domestic epidemic situation, but also the development of the epidemic abroad. At the same time, in the middle and late stages of the epidemic, order restoration works, such as work and production resumption, and study resumption, have attracted widespread attention. The following two reasons may have led to the emergence of the above unique concerns. First, as Chinese citizens were the first nation to experience the coronavirus, they are more sensitive to the development and changes of the epidemic than people in other countries, and second, as the most effective country for epidemic prevention and control, China has the conditions to restore order [31].

It is of paramount importance to promptly guide negative emotions on Sina Weibo. The results of our sentiment analysis show that in 17 topics, the number of topics with positive sentiment values and negative sentiment values are 11 and 6 respectively; for most topics, users have no

Ahead of Print JMIR Publications

obvious positive or negative emotional tendencies. However, it should be noted that there are still some topics where the emotional value of a certain aspect is significant. For example, topics such as pandemic statistics received negative emotional tendencies. The continuous accumulation and spread of negative emotions on Sina Weibo may trigger irrational behavior among citizens, causing users to be affected by rumors or extreme emotions [17], such as group panic and denial of government support; thus, timely guidance on negative emotions is essential.

4.2 Research Implications

With the outbreak and spread of COVID-19 around the world, citizens have turned to social media channels, such as Sina Weibo, to share their opinions, seek clarity and discuss topics related to the crisis. Previous studies have demonstrated that the analysis and control of public behavior and attitudes can effectively help governments cope in times of crises [32]. As many countries start to experience a second wave of COVID-19 outbreaks, social media platforms can collect large amounts of information that reflect public behavior and attitudes. By analyzing this data, we are able to identify the demands and behavioral characteristics of citizens [33].

This paper collected data from Sina Weibo and analyzed it from the point of view of quantity, proportion, emotion and space-time distribution. We identified the degree of attention and the emotional tendency of users to various topics related to COVID-19 and determined the time distribution for each topic and the spatial differences of users' participation in topic discussions. The results obtained allow for deeper understanding of citizen views and attitudes towards COVID-19, which is the premise and basis for the prevention and control of new coronary pneumonia.

4.3 Strengths and Limitations

This paper is based on the existing theories of predecessors, combined with the timely topic of COVID-19. Through analysis of Sina Weibo microblogs, practical conclusions are drawn from the topics related to COVID-19 that are of concern to the public. At the same time, it is undeniable that this article also presents some limitations. First, the microblog posts collected in this study are only some of the blogs related to COVID-19, not all of them. Therefore, the summary of the main concerns of Sina Weibo users is not comprehensive. Secondly, the majority of users are young, so the results of our analysis are more representative of younger citizens. Finally, although the number of Sina Weibo users is relatively large in China, we cannot draw a rushed conclusion that online opinion represents public opinion in general.

Ahead of Print
JMIR Publications

5. Conclusions

The COVID-19 outbreak has caused a crippling impact on the world's economy and exposed numerous challenges to how citizens live and travel. Topics related to COVID-19 have attracted widespread attention and discussion on social media platforms, such as Sina Weibo. By analyzing the microblog posts of COVID-19 topics, we obtained the opinions of citizens on topics related to the virus which can lead to the identification of solutions to solve societal and economic problems [34]. In light of the comparative analysis of the main topics of Sina Weibo and Twitter, the topics that Sina Weibo users concern themselves with show more regional characteristics. Judging from the distribution of microblog posts, the proportion of each topic reveals obvious differences. According to the sentiment analysis results, citizens reveal a positive emotional attitude. However, there still remain some negative emotions among the public, which needs timely guidance from Chinese governments. By analyzing the time distribution of microblog posts, it was observed that the proportion of topics changed significantly during different periods, while the main concerns of users changed with the development of the pandemic. Similarly, by analyzing the publishing location of microblog posts, we identified that in discussion of topics related to COVID-19, user participation revealed obvious regional differences.

References

1. Public Health Emergency Office. The latest news about the COVID-19 by the time of 3rd July [Internet]. 2020. Available from: http://www.nhc.gov.cn/xcs/yqfkdt/202007/8f76a26625044e5391bbd9e8c711b9a7.shtml
2. Johnston KA, Taylor M, Ryan B. Emergency management communication: The paradox of the positive in public communication for preparedness. Public Relations Review [Internet] 2020 Jun [cited 2020 Jul 4];46(2):101903. [doi: 10.1016/j.pubrev.2020.101903]
3. Huang Y, DiStaso M. Responding to a Health Crisis on Facebook: The Effects of Response Timing and Message Appeal. Public Relations Review [Internet] 2020 Sep [cited 2020 Jul 4];46(3):101909. [doi: 10.1016/j.pubrev.2020.101909]
4. Roy KC, Hasan S, Sadri AM, Cebrian M. Understanding the efficiency of social media based crisis communication during hurricane Sandy. International Journal of Information Management [Internet] 2020 Jun [cited 2020 Jul 4];52:102060. [doi: 10.1016/j.ijinfomgt.2019.102060]
5. Chen Q, Min C, Zhang W, Wang G, Ma X, Evans R. Unpacking the black box: How to promote citizen engagement through government social media during the COVID-19 crisis. Computers in Human Behavior [Internet] 2020 Sep [cited 2020 Apr 22];110:106380. [doi: 10.1016/j.chb.2020.106380]
6. Zhu C, Xu X, Zhang W, Chen J, Evans R. How Health Communication via Tik Tok Makes a Difference: A Content Analysis of Tik Tok Accounts Run by Chinese Provincial Health Committees. Int J Environ Res Public Health 2020;13.
7. Yan Q, Wu L, Zheng L. Social network based microblog user behavior analysis. Physica A: Statistical Mechanics and its Applications [Internet] 2013 Apr [cited 2020 Jul 4];392(7):1712–1723.

Ahead of Print JMIR Publications

[doi: 10.1016/j.physa.2012.12.008]
8. Zhang W, Deng Z, Evans R, Xiang F, Ye Q, Zeng R. Social Media Landscape of the Tertiary Referral Hospitals in China: Observational Descriptive Study. Journal of Medical Internet Research [Internet] 2018 Aug 9 [cited 2018 Dec 1];20(8):e249. [doi: 10.2196/jmir.9607]
9. Zhu C, Zeng R, Zhang W, Evans R, He R. Pregnancy-Related Information Seeking and Sharing in the Social Media Era Among Expectant Mothers: Qualitative Study. Journal of Medical Internet Research [Internet] 2019 Dec 4 [cited 2019 Dec 5];21(12):e13694. [doi: 10.2196/13694]
10. Bizid I, Nayef N, Boursier P, Doucet A. Detecting prominent microblog users over crisis events phases. Information Systems [Internet] 2018 Nov [cited 2020 Jul 4];78:173–188. [doi: 10.1016/j.is.2017.12.004]
11. Zhang W, Xu X, Zhang H, Chen Q. Online Participation Chaos: A Case Study of Chinese Government-Initiated E-Polity Square. International Journal of Public Administration [Internet] 2016 Dec 5 [cited 2020 Mar 13];39(14):1195–1202. [doi: 10.1080/01900692.2015.1072218]
12. Lee PSN, So CYK, Lee F, Leung L, Chan M. Social media and political partisanship – A subaltern public sphere's role in democracy. Telematics and Informatics [Internet] 2018 Oct [cited 2020 Jul 4];35(7):1949–1957. [doi: 10.1016/j.tele.2018.06.007]
13. Xie Y, Qiao R, Shao G, Chen H. Research on Chinese social media users' communication behaviors during public emergency events. Telematics and Informatics [Internet] 2017 Jun [cited 2020 Jul 4];34(3):740–754. [doi: 10.1016/j.tele.2016.05.023]
14. Westerman D, Spence PR, Van Der Heide B. Social Media as Information Source: Recency of Updates and Credibility of Information. Journal of Computer-Mediated Communication [Internet] 2014 Jan [cited 2020 Jul 4];19(2):171–183. [doi: 10.1111/jcc4.12041]
15. Chou WS, Hunt YM, Beckjord EB, Moser RP, Hesse BW. Social Media Use in the United States: Implications for Health Communication. Journal of Medical Internet Research [Internet] 2009 Nov 27 [cited 2020 Jul 4];11(4):e48. [doi: 10.2196/jmir.1249]
16. Osatuyi B. Information sharing on social media sites. Computers in Human Behavior [Internet] 2013 Nov [cited 2020 Jul 4];29(6):2622–2631. [doi: 10.1016/j.chb.2013.07.001]
17. Correa T, Hinsley AW, de Zúñiga HG. Who interacts on the Web?: The intersection of users' personality and social media use. Computers in Human Behavior [Internet] 2010 Mar [cited 2020 Jul 4];26(2):247–253. [doi: 10.1016/j.chb.2009.09.003]
18. Leskovec J, Huttenlocher D, Kleinberg J. Signed Networks in Social Media. arXiv:10032424 [physics] [Internet] 2010 Mar 11 [cited 2020 Jul 4]; Available from: http://arxiv.org/abs/1003.2424
19. Zhang W, Deng Z, Hong Z, Evans R, Ma J, Zhang H. Unhappy Patients Are Not Alike: Content Analysis of the Negative Comments from China's Good Doctor Website. J Med Internet Res [Internet] 2018 Jan;20(1):e35. [doi: 10.2196/jmir.8223]
20. Yin J, Karimi S, Lampert A, Cameron M, Robinson B, Power R. Using Social Media to Enhance Emergency Situation Awareness: Extended Abstract. Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence 2015. p. 5.
21. Simon T, Goldberg A, Adini B. Socializing in emergencies—A review of the use of social media in emergency situations. International Journal of Information Management [Internet] 2015 Oct [cited 2020 Jul 4];35(5):609–619. [doi: 10.1016/j.ijinfomgt.2015.07.001]
22. Hajjem M, Latiri C. Combining IR and LDA Topic Modeling for Filtering Microblogs. Procedia Computer Science [Internet] 2017 [cited 2020 Jul 4];112:761–770. [doi: 10.1016/j.procs.2017.08.166]
23. Rani R, Lobiyal DK. Performance evaluation of text-mining models with Hindi stopwords lists. Journal of King Saud University - Computer and Information Sciences [Internet] 2020 Mar [cited 2020 Jul 4]; [doi: 10.1016/j.jksuci.2020.03.003]
24. Blei DM, Ng AY, Jordan MI. Latent Dirichlet Allocation. Journal of Machine Learning Research [Internet] 2003 [cited 2020 Oct 6];3(Jan):993–1022. Available from: https://www.jmlr.org/papers/v3/blei03a

25. Arora S, Ge R, Halpern Y, Mimno D, Moitra A, Sontag D, Wu Y, Zhu M. A Practical Algorithm for Topic Modeling with Provable Guarantees. :9.
26. Pavlinek M, Podgorelec V. Text classification method based on self-training and LDA topic models. Expert Systems with Applications [Internet] 2017 Sep [cited 2020 Jul 4];80:83–93. [doi: 10.1016/j.eswa.2017.03.020]
27. Stevens K, Kegelmeyer P, Andrzejewski D, Buttler D. Exploring Topic Coherence over Many Models and Many Topics. Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning [Internet] Jeju Island, Korea: Association for Computational Linguistics; 2012 [cited 2020 Oct 6]. p. 952–961. Available from: https://www.aclweb.org/anthology/D12-1087
28. Amoualian H, Lu W, Gaussier E, Balikas G, Amini MR, Clausel M. Topical Coherence in LDA-based Models through Induced Segmentation. Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) [Internet] Vancouver, Canada: Association for Computational Linguistics; 2017 [cited 2020 Aug 18]. p. 1799–1809. [doi: 10.18653/v1/P17-1165]
29. Abd-Alrazaq A, Alhuwail D, Househ M, Hamdi M, Shah Z. Top Concerns of Tweeters During the COVID-19 Pandemic: Infoveillance Study. Journal of Medical Internet Research [Internet] 2020 Apr 21 [cited 2020 May 13];22(4):e19016. [doi: 10.2196/19016]
30. The State Council Information Office of the People's Republic of China. Fighting Covid-19 China in Action [Internet]. 2020 Jun. Available from: https://covid-19.chinadaily.com.cn/a/202006/08/WS5edd8bd6a3108348172515ec.html
31. Shaoyong Chen, Huanming Zhang, Min Lin, Shuanghuan Lv. Comparision of microblogging service between Sina Weibo and Twitter. Proceedings of 2011 International Conference on Computer Science and Network Technology [Internet] Harbin, China: IEEE; 2011 [cited 2020 Jul 4]. p. 2259–2263. [doi: 10.1109/ICCSNT.2011.6182424]
32. Xiaomei Z, Jing Y, Jianpei Z, Hongyu H. Microblog sentiment analysis with weak dependency connections. Knowledge-Based Systems [Internet] 2018 Feb [cited 2020 Jul 4];142:170–180. [doi: 10.1016/j.knosys.2017.11.035]
33. Fersini E, Pozzi FA, Messina E. Approval network: a novel approach for sentiment analysis in social networks. World Wide Web [Internet] 2017 Jul [cited 2020 Jul 4];20(4):831–854. [doi: 10.1007/s11280-016-0419-8]
34. Saroj A, Pal S. Use of social media in crisis management: A survey. International Journal of Disaster Risk Reduction [Internet] 2020 Sep [cited 2020 Jul 4];48:101584. [doi: 10.1016/j.ijdrr.2020.101584]

Supplementary Files

Ahead of Print
JMIR Publications

Multimedia Appendixes

Examples of microblogs for each topic.
URL: http://asset.jmir.pub/assets/7cb8a5831c37682b6866011a21e4a1d0.docx

Topics in Chinese.
URL: http://asset.jmir.pub/assets/814fe0a4c0ecbdb2b048c28af2eff756.docx

Ahead of Print
JMIR Publications