On the Development of an Information System for Monitoring User Opinion and its Role for the Public

Vladislav Karyukin (vladislav.karyukin@kaznu.kz)
Al-Farabi Kazakh National University

Galimkair Mutanov
Al-Farabi Kazakh National University

Zhanl Mamykova
Al-Farabi Kazakh National University

Gulnar Nassimova
Al-Farabi Kazakh National University

Saule Torekul
Al-Farabi Kazakh National University

Zhanerke Sundetova
Al-Farabi Kazakh National University

Matteo Negri
Fondazione Bruno Kessler

Research Article

Keywords: Social media, social networks, social mood, sentiment analysis, analytics platforms, machine learning, OMSystem, coronavirus disease, vaccination

Posted Date: December 1st, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1122375/v1

Abstract

Social media services and analytics platforms are rapidly growing. A large number of various events happen mostly every day, and the role of social media monitoring tools is also increasing. Social networks are widely used for managing and promoting brands and different services. Thus, most popular social analytics platforms aim for business purposes while monitoring various social, economic, and political problems remains underrepresented and not covered by thorough research. Moreover, most of them focus on resource-rich languages such as the English language, whereas texts and comments in other low-resource languages such as the Russian and Kazakh languages in social media are not represented well enough. So, this work is devoted to developing and applying the information system called the OMSystem for analyzing users' opinions on news portals, blogs, and social networks in Kazakhstan. The system uses sentiment dictionaries of the Russian and Kazakh languages and machine learning algorithms to determine the sentiment of social media texts. The whole structure and functionalities of the system are also presented. In the experimental part, the system's monitoring of the healthcare, political and social aspects of the most relevant topics connected with the vaccination against the coronavirus disease are thoroughly observed and analyzed. The analysis allowed discovering the public social mood in the cities of Almaty and Nur-Sultan and other large regional cities of Kazakhstan. The system's study included two extensive periods: 10-01-2021 to 30-05-2021 and 01-07-2021 to 12-08-2021. In the obtained results, people's mood and attitude to the Government's policies and actions were studied by such social network indicators as the level of topic discussion activity in society, the level of interest in the topic in society, and the mood level of society. These indicators calculated by the OMSystem allowed careful identification of alarming factors of the public (negative attitude to the government regulations, vaccination policies, trust to vaccination, etc.) and assessment of the social mood.

1. Introduction

The rapid development of the Internet, social networks, online services, and other web resources initiated a great interest in the use of information from social networks and the great online activity of users. New events occur almost daily, and their relevance is constantly changing. In this regard, the technologies of "monitoring social networks" (social listening) and content analysis are gaining great popularity.

In many cases, social networks are used to solve a wide range of business tasks: managing and promoting brands [1], advertising goods and services, creating distribution channels for goods, etc. In addition to business tasks [2], there is a great need for monitoring social networks [3] and content analysis in other areas. Critical topics in politics [4], economics [5], healthcare, medicine, culture, and other areas are gaining great popularity in the media space [6]. Despite this fact, most popular social analytics platforms still focus on business purposes leaving significant social, economic, and political problems uncovered. Moreover, most of them focus on resource-rich languages such as the English language, whereas texts and comments in other low-resource languages such as the Russian and Kazakh languages are underrepresented.

Thereby, a new opinion monitoring information system, the OMSystem, which pays much attention to the political, economic, healthcare, education, culture, ecology, and civil society topics, has been developed. This multifunctional platform monitors the media space of Kazakhstan and supports the Kazakh and Russian languages, which allows analyzing the media space efficiently. The OMSystem supports Kazakhstan's leading news portals and important popular social networks like Facebook, VKontakte, Instagram, Twitter, and YouTube. The core part of the system is the evaluation of the public's mood and "social well-being" with the use of the sentiment analysis tool and the social mood indicators such as the level of topic discussion activity in society, the level of interest in the topic in society, and the level of social mood. The sentiment analysis tool determines the sentiment [7] of the public mood, the range of interests, and information dissemination. It also identifies current problematic issues in society and tracks the dynamics of user involvement in a certain topic. This tool uses the sentiment analysis (SA) methods generally presented by three main approaches: lexicon-based, machine learning-based, and deep learning-based.

This paper describes the architecture of the OMSystem, main modules, and functionalities of this platform, focusing on the module of defining the social mood of society. The experimental part is devoted to the topic of vaccination against coronavirus infection. Many scientific articles review this topic, and research in this field is especially demanded today. The experimental

results have been thoroughly analyzed, providing exciting findings of the public's attitude to the vaccination campaign, vaccination policies, and the Government's activities and methods of combating the pandemic. The reasons for people's negative moods on this topic have also been extensively analyzed.

The rest of the paper is organized in the following way: Section 2 provides an overview of the related works to this paper. Section 3 describes the features of popular social analytics platforms for brand monitoring, highlighting the essential missing tools implemented in the OMSystem. Sections 4 and 5 describe the structure, functionalities, and module for social mood evaluation. Section 6 describes and discusses the experiment about the public's attitude towards the vaccination against coronavirus infection. Finally, in Section 7, we summarize all the previously described sections, analyze the obtained results, and outline directions for future research.

2. Related Works

In recent years, the active development of web technologies has made it possible to analyze users' moods on various topics. At the same time, marketing campaigns interested in learning users' opinions and developing many strategies for increasing the flow of customers and profits play a significant role in data analytics. The manual search and filtering of users' views on websites remain challenging because of their vast number. Therefore, special tools have been developed to automatically track, summarize, and visualize information from social content to solve this problem. In [8], SA of the popular smartphone brand was presented. Data was collected from Twitter using a web crawler that searches through particular hashtags. [9] demonstrates an open framework for monitoring, analyzing, and receiving media content. This framework allows you to collect, index, and retrieve data using the Representational state transfer application programming interface (REST API) from the following sources: Twitter, Facebook, YouTube, Google +, and Flickr. [10] presents an analysis of the statements of many political leaders, diplomats, journalists, and other media figures on the Twitter platform, the most active social network covering these issues. [11] shows an architecture that combines SA and community discovery to understand trends, approaches, business, and policy views on topics such as shopping, politics, Covid-19, and electric vehicles. At the same time, many works are devoted to describing analytics platforms and social networks and processing of texts for SA. [12] describes the steps of preprocessing, vectorization, and classification of the textual data using machine learning (ML) algorithms. [13] pays great attention to studying the critical approaches of the most efficient ML algorithms for SA. That work showed that the support vector machine (SVM) and naïve Bayes (NB) classifier are more effective than other algorithms. The classification of Twitter posts is also performed in [14], where the primary role is assigned to the k-nearest neighbours (k-NN) and SVM. [15] provides detailed SA of user opinions from Twitter and Facebook social networks using convolutional neural networks (CNN), recurrent neural networks (RNN), long short-term memory (LSTM) neural network, and hybrid approaches. In [16], comments on controversial political discussions in German on YouTube were conducted. SA was performed with various word embeddings, ML algorithms, and RNN. Then the classification efficiency was assessed using the following metrics: Precision, Recall, and F1-score. A new and more advanced approach to text classification using one CNN and two LSTM layers was described in [17].

All these works were mainly devoted to the analysis of texts in the English language. However, most texts and user comments are written in the Russian and Kazakh languages in the Kazakh media space. Thus, it became necessary to specifically analyze the works dealing with these languages. The sentiment classification of Russian tweets using logistic regression (LR), XGBoost, and CNN was carried out in [18, 19]. Unfortunately, the works devoted to the SA of Kazakh texts are greatly underrepresented. [20, 21] implemented only a dictionary approach formalizing rules for defining the sentiment of phrases in texts. ML and NN approaches had a limited reflection in these works. In addition, they neither described any open-source analytics platforms nor provided functionalities for evaluating society's SA and social mood in the Russian and Kazakh media spaces. Thus, various foreign and Kazakh analytics platforms were thoroughly investigated in the next section.

3. Analytics Platforms

The widespread development of Internet technologies, social networks [22], and data analytics has led to numerous tools and analytics platforms for promoting the brand, monitoring public opinions, and assessing social well-being as one of the main tools for determining the socio-economic system in the context of sustainable advancement.

Currently, the foreign market is represented by many tools for monitoring social networks [23], content analysis, and brand promotion. Therefore, the marketers distinguished a list of the most popular and advanced analytics platforms: Sprout social, Hubspot, Buzzsumo, Hootsuite, Brandmention, IQBuzz, and Snaplytics take an essential place.

Sproutsocial [24] is a multifunctional analytics tool that allows comparing results in several networks efficiently. It is very useful when it is required to count links on Twitter [25], measure the growth of Instagram followers, evaluate participation in LinkedIn, and much more. This tool then provides an opportunity to evaluate results using understandable visualized reports. Sprout social includes marketing, social media management, and analytics of various leading brands and agencies, including Chipotle, Subaru, Zendesk, etc.

Hubspot [26] is a tool that allows marketers to obtain comparative information about the level of engagement on social networks and reflect on past efforts made to support high customer interest in their products. HubSpot provides a detailed overview of how social media affects profit margins and enables you to report on collected data quickly and efficiently. At the same time, it gives an opportunity to compare different platforms, track and view brands on social networks, and understand how the target audience watches business content. Another essential feature of the HubSpot analytics tool is the ability to analyze indicators specific to social networks and the entire path of the client. This tool also provides information about marketing tactics that are most effective for businesses and their impact on social media campaigns.

BuzzSumo [27] is an excellent resource for analyzing the social interaction of any particular content. The tool allows searching for information based on requests on the Internet, taking into account various factors, including likes and reposts. The advanced search engine of BuzzSumo finds the most relevant content by topic, author, and domain. The service prompts which directions most respond to the initially selected audience. Trying to choose the most accurate direction of content creation, it receives valuable information about answers on social networks. In addition, this tool allows collecting statistics of the number of reposts of a certain message on a blog on such social networks as Facebook [28], Twitter, and Pinterest. A feature of the tool is the ability to track the effectiveness of competitors as part of a content marketing campaign. BuzzSumo also easily determines competitors' activity in social networks and identifies key people in a particular area. Such an analysis can help to see which posts receive the most engagement and use this data to adjust the content strategy.

Hootsuite [29] is one of the most popular multifunctional services for working on social networks. The emphasis in this service is on working with Twitter, and, first of all, Hootsuite will be useful for those who maintain several accounts at once. Hootsuite also works successfully with Facebook, LinkedIn, MySpace, and Foursquare accounts and blogs on WordPress. HootSuite offers a wide range of analytical capabilities, such as connecting Google Analytics on the site and viewing graphs for comparing the number of tweets and the popularity of links. Hootsuite additionally allows you to post on all social networks on a specific schedule. The tool also allows you to track recent social trends and brand mentions.

Brandmention [30] is one of the most powerful platforms for free search and analysis of social networks. The system also offers SA, related keywords, popular sources, etc. Brandmention searches over 100 social networks [31], including social bookmarks, blogs, forums, social services, and more. In addition, data can be exported or configured for e-mail.

IQBuzz [32] is a professional tool for analyzing and managing reputation on the Internet and a social network monitoring service [33]. IQBuzz tracks many sources and platforms such as Twitter, Yandex, LiveInternet, LiveJournal, various blogs, video hosting services such as RuTube and YouTube, various news, entertainment, and specialized services, thematic and regional portals. One of the key advantages of the service is the ability to connect new sources and Internet resources for monitoring.

Snaplytics [34] is a platform that analyzes Snapchat and Instagram stories. Today, millions of active Snapchat and Instagram users present stories as an excellent method of promotion on Instagram. This application also allows you to see peaks and slumps of views.

In Kazakhstan, there are only a few brands and social analytics platforms. Their research is mostly restricted to SA with the use of dictionary and ML approaches [20-21]. Among the most advanced applications are the **iMAS** [35] and the **Alem Media Monitoring** [36], which work with the Russian and Kazakh languages. The iMAS platform provides SA on specified topics for a

given period. The Alem Media Monitoring is software designed to analyze public opinion in the Internet space. This system allows collecting information on certain topics from news portals and social networks [37], determining the sentiment of texts using ML algorithms, visualizing all the performed analyses, and compiling and uploading reports. Unfortunately, these platforms are not open-source, and the information provided on their official websites demonstrates the study by three sentiment classes (positive, negative, and neutral) of texts and comments, the sources (news portals, social networks, and blogs), and periods of monitoring, visualizing them with different graphics and making reports in the word, excel and pdf formats. Nevertheless, there is no description of how these systems estimate the public's social mood. Moreover, the research papers devoted to the iMAS and the Alem Media Monitoring platforms have not been found online. The proposed OMSystem was first described in [38]. The following sections demonstrate the structure, functionalities, and module for evaluating the social mood of society.

4. The Omsystem Information System Design Methodology

The OMSystem, the first automatic tool developed to analyze the opinions of Kazakhstani users expressed through news portals, blogs, and social networks, was developed to provide the complex analysis of the public's social mood and cover the parts skipped in other analytics platforms in Kazakhstan. The OMSystem allows monitoring web resources and social networks with subsystems for modelling "social well-being" [39] and supporting sentiment dictionaries of the Russian and Kazakh languages and ML algorithms for determining the sentiment of texts and user comments. The OMSystem supports Kazakhstan's leading news portals and popular social networks like Facebook, VKontakte, Instagram, Twitter, and YouTube. The platform's main tasks are the operational monitoring of the information space and social networks on the most important topics in society. They unambiguously determine the scale of the problem, public opinion, and their quick explanation, analyze the dynamics of the commercial brand, events, and references to activities, and, in turn, assess the degree of "social well-being."

This system allows working with texts in the Kazakh and Russian languages. It also has built-in modules for connecting to the application programming interfaces (APIs) of social networks: Vkontakte [40], Facebook [41, 42], Twitter [43, 44], Instagram [45, 46], YouTube [47], Telegram [48], and Odnoklassniki [49]. The OMSystem automatically determines the language of the text (Russian, Kazakh, smiles/characters) and the sentiment of the topic, as negative, positive, or neutral, using a sentiment dictionary and ML algorithms. Furthermore, there is a possibility to record the time range in the system when monitoring social networks (for a year, for six months, for three months, for a month, for a week, for a day, etc.). The OMSystem also allows building visual reports on the monitoring results in various graphs and charts (pie, histogram, chart, graph, and others). At the same time, the platform provides ways to identify the profile of a social network participant by reading profile data and counting the activity of a participant in a topic by a number of comments, likes, and reposts.

The OMSystem architecture schematically shown in Figure 1 includes the following components:

- Data sources: They are represented by news portals, blogs, and social networks.
- Connector module: It is used for the connection to data sources.
- The linguistic constructor module: It is used for creating sentiment dictionaries that include words belonging to any of the three classes: positive, negative, and neutral.
- Data analysis and processing module: It uses sentiment dictionaries and ML algorithms for sentiment analysis. In addition, this module creates social analytics defining social mood.
- Results module: It contains a formed relational database of texts and comments, analytical reports, graphics and tables.

The sentiment analysis tool, labelling texts and user comments in three sentiment classes (positive, neutral, and negative), is the core part of the OMSystem. The sentiment classes are assigned with the use of the hybrid approach: the lexicon-based (sentiment dictionaries) and the ML-based. The lexicon-based approach assigns a label by the largest number of words of one of three sentiment classes. The ML-based approach uses the trained ML models with the highest efficiency in terms of accuracy, precision, recall, and F1-score such as NB, LR, SVM, k-NN, decision tree (DT), random forest (RF), and XGBoost.

5. Defining The Social Mood Of Society

While OMSystem provides a comprehensive analysis of the texts of Kazakhstan Internet resources, reveals the sentiment of user opinions using ML methods, it also allows evaluating the semantic profile of society's response to various events. The models of engagement assessment standards are considered to implement these steps. They are based on the method of measuring social network indicators for social media marketing management (SMMM) with the use of special SocialBakers formulas from Facebook [50]. The presented metrics are considered and adapted for social analytics. They are presented below:

- the level of interest in the topic in society (R_{cr});
- the level of topic discussion activity in society (R_{CE});
- the level of social mood (R_{rs}).

The level of interest in the topic is calculated using the following formula:

$$R_{CT} = \frac{CT \times 100\%}{\max_{CT}}, \qquad (1)$$

where CT is the number of texts or comments found on a particular topic. $\max_{ct}$is the maximum number of texts or comments on a certain topic (set by the expert for a certain time). The range of values starts from 0% and is not bounded. If the value exceeds 100%, it means that this topic is of a great interest.

R_{CE}determines interaction in social networks and shows the level of topic discussion activity in society. This indicator allows assessing how differently the audience reacts to the categories of events in society. It is calculated using the formula:

$$R_{CE} = \frac{\frac{L+R+C}{CP}}{CS} \times 100\%, \qquad (2)$$

where CS is the sum of the number of subscribers, CP is the number of texts found on a certain topic, C is the number of comments, L is the number of likes, R is the number of reposts. The range of values starts from 0% and is not bounded. As there are many topics on each news portal or a group in a social network and all users and subscribers cannot discuss them all, the level of topic discussion activity is usually not a big number.

R_{rs} is the level of social mood, which is defined by the maximum value of the sums of positive, neutral, and negative texts or comments on a certain topic.

6. Experimental Part

A relevant topic of vaccination against coronavirus infection [51] is taken for analysis in the experimental part. This topic is very important due to the active vaccination [52] of people in the world and Kazakhstan. A large number of news articles have been written on this topic, and users actively comment on various issues related to it. The opinions of users stand out with positive, neutral, and negative sentiments. The experimental part chooses a list of keywords and phrases in the Russian language to monitor the corresponding topics. In the following description of the experiment, all words and phrases originally in the Russian language are translated into the English language for convenience and the right understanding. These keywords and phrases are "Vaccination in Kazakhstan," Covid [53], Coronavirus [54], Sputnik, "Russian vaccine," Pfizer [55], QazVac, Hayat, Sinovac, Sinopharm [56], "Vaccine rejection," "Fear of vaccination," "Choice of the vaccine," "Vaccine effectiveness," "Lack of confidence in the vaccine," and Tsoi (the last name of the Minister of Health of the Republic of Kazakhstan).

The steps of the OMSystem's functionalities [39] are presented in Figure 2.

When the web-crawler of the OMSystem is launched, it parses texts and user comments from the specified list of sources (Kazakhstan's news portals, social networks, and blogs). The parsed texts are aggregated in the designated PostgreSQL database. Then the following steps are applied to texts:

- Text preprocessing
- Stemming
- Vectorization
- Assignment of sentiment labels

In the preprocessing step, all words are transformed to the lowercase register. Then punctuation marks, digits, and other special symbols that do not carry any significant meaning are removed. Additionally, it is required to delete frequent words (i.e., stop words such as 'and,' 'or,' 'in,' 'on,' 'at,' 'for,' etc.), which do not bring any significant meaning [38]. However, 'to be' and 'is' stop words are left because they are met in such expressions as "to be vaccinated," "is vaccinated," and others and they are important for the analyzed topic.

The stemming step reduces the number of words with similar meanings by eliminating affixes and endings to gain their roots. Russian words are processed by 'SnowballStemmer' from the Python NLTK library. The text vectorization step transforms texts into a numeric vector representation on which ML algorithms are applied [38]. The vectorization is done with the use of the TF-IDF metric that considers the importance of words in the text. After the texts are vectorized, the trained ML models are applied to label them in three sentiment classes.

Next, the number of words in texts and comments is counted, and the most frequently used ones are displayed in pivot tables. The OMSystem [38-39] performs calculations for two periods: 10–01–2021 to 30–05–2021 (Table 1) and 01–07–2021 to 12–08–2021 (Table 3) and two groups of cities: Almaty (the largest city of Kazakhstan) and Nur-Sultan (the capital of Kazakhstan), and large regional cities of Kazakhstan. The reason for choosing these periods is since a vaccination program was launched at that time, and it is possible to estimate the level of interest in this topic in society.

Table 1. Analysis by topics for period 1

Resource Set	News portals, Vkontakte, Facebook, Instagram, Youtube			Vkontakte, Facebook, Instagram, Youtube			
Search Period:	**from "10-01-2021" to "30-05-2021"**						
Location:	**Cities of Almaty and Nur-Sultan**			**Large regional cities of Kazakhstan**			
Number of results (texts + comments)	~19340			~1228			
Number of texts	~4919			~122			
Number of comments	~14421			~1106			
The level of social mood by results	**Positive**	**8944**		Positive	396		
	Negative	8152		**Negative**	**683**		
	Neutral	1082		Neutral	66		
	Undefined	1162		Undefined	83		
The level of social mood by texts	**Positive**	**3829**		**Positive**	**56**		
	Negative	960		Negative	43		
	Neutral	123		Neutral	11		
	Undefined	7		Undefined	12		
The level of social mood by comments	Positive	5115		Positive	340		
	Negative	**7192**		**Negative**	**640**		
	Neutral	959		Neutral	55		
	Undefined	1155		Undefined	71		
The level of topic discussion activity in society	**~0.48%**			**~0.08%**			
The level of interest in the topic in society	**~491%**			**~12.2%**			
Engagement level				**Engagement level**			
Views	~9M			~341K			
Comments	~14K			~1K			
Reposts	~2K			~249			
Likes	~32K			~2K			
Dislikes	~2K			~305			
Total Engagement Level	**~9M**			**~345K**			
Popular words				**Popular words**			
by texts		**by comments**		**by texts**		**by comments**	
Word	**Frequency of use**	**Word**	**Frequency of use**	**Word**	**Frequency of use**	**Word**	**Frequency of use**
Coronavirus	2374 (3.51%)	To be	1598 (1.38%)	Coronavirus	118 (1.29%)	To be	148 (1.52%)
Kazakhstan	1811	**Vaccine**	**1112**	To be	113	Person	138

	(2.68%)		**(0.96%)**		(1.24%)		(1.42%)
Vaccine	824 (1.22%)	Person	1097 (0.94%)	Area	112 (1.23%)	**Vaccine**	**105 (1.08%)**
Person	653 (0.96%)	Can	564 (0.48%)	Kazakhstan	101 (1.11%)	People	92 (0.94%)
Covid-19	540 (0.80%)	Is	532 (0.46%)	Vaccine	68 (0.74%)	Kazakhstan	62 (0.63%)
Day	526 (0.77%)	Kazakhstan	477 (0.41%)	Aktyubinsk	59 (0.64%)	Year	53 (0.54%)
Vaccination	524 (0.77%)	**Necessary**	**472 (0.40%)**	Year	55 (0.60%)	**Necessary**	**44 (0.45%)**
News	502 (0.74%)	Year	432 (0.37%)	Person	53 (0.58%)	Virus	42 (0.43%)
Almaty	443 (0.65%)	People	415 (0.35%)	Vaccination	52 (0.57%)	Country	41 (0.42%)
New	433 (0.64%)	Virus	370 (0.32%)	Tenge	49 (0.53%)	Can	40 (0.41%)
Country	427 (0.63%)	To speak	337 (0.29%)	Reference	48 (0.52%)	**Vaccination**	**40 (0.41%)**
To be	394 (0.58%)	Country	327 (0.28%)	Zone	46 (0.50%)	Power	35 (0.36%)
Case	371 (0.54%)	**To do**	**327 (0.28%)**	Case	42 (0.46%)	Good	29 (0.29%)
The first	358 (0.53%)	**Vaccination**	**325 (0.28%)**	To attach	40 (0.43%)	Russia	29 (0.29%)
Area	325 (0.48%)	To tell	317 (0.27%)	Later	39 (0.42%)	Simply	29 (0.29%)
Zone	310 (0.45%)	To want	312 (0.27%)	Child	36 (0.39%)	Covid	29 (0.29%)
Ministry of Health	282 (0.41%)	Nobility	297 (0.25%)	Can	34 (0.37%)	Child	28 (0.28%)
To reveal	281 (0.41%)	Money	286 (0.24%)	Doctor	33 (0.36%)	Inoculation	28 (0.28%)
Tsoi	274 (0.40%)	Another	281 (0.24%)	Healthcare	32 (0.35%)	World	27 (0.27%)
Year	271 (0.40%)	Good	279 (0.24%)	Region	31 (0.34%)	Quarantine	27 (0.27%)

The sentiment charts of the first period for the cities of Almaty and Nur-Sultan and large regional cities are shown in Figure 3.

Based on the results of the analysis of Table 1, it is possible to evaluate the content of texts and comments, taking into account the list of the most popular words. Furthermore, looking at the analysis of popular words in the context of regional cities, we will see that they coincide with the content in the cities of Almaty and Nur-Sultan. According to the obtained results, the level of interest in this topic is significantly higher in the cities of Almaty and Nur-Sultan (491%) than in other large regional cities (12.2%). In addition, the level of topic discussion is also higher in the two main cities of the country (0.48%) than in other ones (0.08%). The level of social mood of texts and comments differs significantly, with the positive sentiment prevailing over the negative sentiment in texts and the negative sentiment prevailing over the positive sentiment in comments. It shows that texts in

social media positively cover the topic of vaccination, while people's attitude is the opposite. After the system had created a summary analysis, the gained texts and comments were manually read and investigated. Their examples are presented in Table 2. The public showed a very negative reaction to all governmental measures related to the vaccination campaign in the winter and spring seasons.

Table 2. Texts and comments for period 1

№	Date	Sentiment	Text	Sentiment	Comments
1	19–01–2021	positive	The Head of the Government instructed the Ministry of Health of the Republic of Kazakhstan to ensure the readiness of medical organizations for the start of the mass vaccination of the population with the “Sputnik V” vaccine from February 1.	negative	We do not need your vaccine; go to poison others with these chemicals
				negative	Look for idiots elsewhere. Madhouse
				negative	Experiments on humans are like this, especially when all sane scientists deny the vaccine’s effectiveness. So it is time to be vaccinated!
				negative	Madness! Even in Russia, they did not really test it. Did they decide to test it on the Kazakhs? What a madhouse?
				positive	Ready to become a test subject for a fee. Where to go?
2	30–01–2021	positive	First of all, vaccines against COVID-19 will be sent to health workers in eight regions of Kazakhstan, in which there is a high incidence of coronavirus. On Monday, February 1, vaccination against coronavirus starts, but vaccines have not yet been brought to the Aktobe region.	negative	Now we will look before the vaccine and after.
				positive	Soon, vaccinations will start everywhere, and it will be much more difficult for the coronavirus to spread. So the epidemic will end.
				negative	Why do not we start with the deputies?
				positive	The entire Government with their families must be at the forefront. May they get the best,

					we undoubtably agree this time
3	08–05–2021	positive	In Kazakhstan, 34% of residents have changed their attitude towards vaccination against COVID-19 for the better. At the same time, 23% are skeptical against vaccinations. 9% have recently changed their minds in a negative direction. 4% do not trust the vaccine, 3% do not dare to get vaccinated, as they recently got sick. Who do you belong to?	negative	The data is not exact. More than half of the population of Kazakhstan do not believe
				negative	Where do these statistics come from? For example, no one asked me)
				negative	I did it because I wanted to save my family. The opinions of others do not matter, but my health and safety do.
				neutral	There is no way
				negative	Suicidal people are getting vaccinated
4	01–02–2021	positive	Kazakhstanis are intended to be classified by color in terms of whether they passed the polymerase chain reaction (PCR) test and what the result was, zakon.kz reports. According to the press service of the Republic of Kazakhstan, the data will be reflected in the “Ashyq” application developed by the Ministry of Digital Development, Innovation and Aerospace Industry jointly with the Ministry of Health of the Republic of Kazakhstan.	negative	I am crazy with all sorts of this bullshit to torture the people
				negative	Well! It is straight racism: yellow and red. I disagree
				negative	It is a total control under the guise of coronavirus
				negative	"Divide and conquer" is a working scheme from the ancient time
				negative	Scumbugs! I knew it would come to this!
5	26–01–2021	positive	Mass vaccination of the population against COVID-19 will begin in Kazakhstan on February 1, Kazakh Health Minister Alexei Tsoi said during a government meeting on Tuesday. It is planned to vaccinate up to six million people by the end of the year	negative	Are we their guinea pigs or what? Go away. Check your vaccine to the end first, then to the people.
				negative	It is necessary to start with the ministers and deputies. Whoever survives will remain in office, who does not

		survive, and to hell with them!
	negative	They want to test the effectiveness of the vaccine on us
	negative	Let them first try this vaccine on themselves. We did not invent this infection; it was not for us to die for it.
	positive	Yes. Nevermind. As they said, it will be “finally,” and vaccination is already in full swing

Table 3. Analysis by topics for period 2

Resource Set	News portals, Vkontakte, Facebook, Instagram, Youtube		Vkontakte, Facebook, Instagram, Youtube	
Search Period:	**from “07/01/2021” to “08/12/2021”**			
Location:	**Cities of Almaty and Nur-Sultan**		**Large regional cities of Kazakhstan**	
Number of results (texts + comments)	~2133		~157	
Number of texts	~1285		~69	
Number of comments	~848		~88	
The level of social mood by results	Positive	1029	Positive	52
	Negative	955	Negative	58
	Neutral	80	Neutral	10
	Undefined	69	Undefined	37
The level of social mood by texts	Positive	739	Positive	21
	Negative	544	Negative	18
	Neutral	1	Neutral	5
	Undefined	1	Undefined	25
The level of social mood by comments	Positive	290	Positive	31
	Negative	411	Negative	40
	Neutral	79	Neutral	5
	Undefined	68	Undefined	12
The level of topic discussion activity in society	~0.01%		~0.03%	
The level of interest in the topic in society	~128%		~6.9%	
Engagement level			**Engagement level**	
Views	~34K		~42K	
Comments	~848		~97	
Reposts	~825		~46	
Likes	~2K		~123	
Dislikes	~35		~0	
Total Engagement Level	**~38K**		**~42K**	

Popular words				Popular words			
by texts		by comments		by texts		by comments	
Word	Frequency of consumption	Word	Frequency of consumption	Word	Frequency of consumption	Word	Frequency of consumption
To be	1786 (1.00%)	**Vaccine**	**143 (1.59%)**	Coronavirus	52 (2.15%)	Person	11 (1.59%)
Kazakhstan	1630 (0.91%)	Person	82 (0.91%)	Reference	44 (1.82%)	**Vaccine**	**11 (1.59%)**
Person	1493 (0.83%)	To be	63 (0.70%)	To attach	40 (1.66%)	To be	11 (1.59%)
Year	1268 (0.71%)	**Vaccination**	**46 (0.51%)**	Area	36 (1.49%)	Simply	8 (1.16%)
Coronavirus	1213 (0.68%)	Kazakhstan	39 (0.43%)	Strain	22 (0.91%)	People	6 (0.87%)
Vaccination	**1201 (0.67%)**	Later	38 (0.42%)	Pavlodar	22 (0.91%)	Though	4 (0.58%)
Vaccine	**1018 (0.57%)**	Child	33 (0.36%)	Kazakhstan	21 (0.87%)	Level	4 (0.58%)
Case	808 (0.45%)	Can	33 (0.36%)	Heading	20 (0.83%)	To buy	4 (0.58%)
Country	750 (0.42%)	To speak	33 (0.36%)	**Vaccine**	**20 (0.83%)**	Proper	4 (0.58%)
Infection	724 (0.40%)	Necessary	30 (0.33%)	Url	20 (0.83%)	Virus	4 (0.58%)
Covid-19	722 (0.40%)	To know	30 (0.33%)	To be	19 (0.78%)	Guilty	4 (0.58%)
Can	692 (0.38%)	Is	30 (0.33%)	Year	17 (0.70%)	Strain	3 (0.43%)
Area	674 (0.37%)	Covid	29 (0.32%)	Person	16 (0.66%)	**To make**	**3 (0.43%)**
July	638 (0.35%)	People	27 (0.30%)	Can	14 (0.58%)	**Inoculation**	**3 (0.43%)**
More	635 (0.35%)	**To do**	**24 (0.26%)**	**Vaccination**	**14 (0.58%)**	In a row	3 (0.43%)
Day	633 (0.35%)	Year	24 (0.26%)	Health care	13 (0.53%)	Life	3 (0.43%)
Work	631 (0.35%)	Doctor	24 (0.26%)	Pavlodar	12 (0.49%)	Any	3 (0.43%)
New	630 (0.35%)	**To be ill**	**24 (0.26%)**	Doctor	12 (0.49%)	Small	3 (0.43%)
Coronavirus	612 (0.34%)	To tell	23 (0.25%)	To work	11 (0.45%)	To know	3 (0.43%)
Patient	603 (0.33%)	Virus	21 (0.23%)	To become	10 (0.41%)	Delta	3 (0.43%)

The sentiment charts of the second period for the cities of Almaty and Nur-Sultan and large regional cities are shown in Figure 4.

The analysis of Table 3 suggests that there remains a high level of public interest in the topic during the summer. The level of interest in this topic is higher in the cities of Almaty and Nur-Sultan (128%) than in large regional cities (6.9%). The level of topic discussion activity is lower than in period 1. It is caused by fewer comments on the considered topics during a shorter time of monitoring. The following values are gained in the context of cities: 0.01% for Almaty and Nur-Sultan and 0.03% for the large regional cities. The level of social mood of texts and comments shows a situation similar to period 1. This period's obtained texts and comments were also manually analyzed to reveal interesting points. It is noted that texts cover the planned children's vaccination topics, the appearance of new strains of coronavirus, the increase in the number of cases of unvaccinated people's

disease, and the supply of a new Chinese vaccine to the country. The corresponding examples of the texts and comments are presented in Table 4.

Table 4. Texts and comments for period 2

№	Date	Sentiment	Text	Sentiment	Comments
1	21–07–2021	positive	It is planned to start vaccination of children against coronavirus in Kazakhstan at the end of this year. What do they plan to vaccinate with, and will vaccination be voluntary?	negative	People stand up to protect children. Healthcare is not able to protect children from vaccine refusal
				positive	In the USA, all children are vaccinated. If we can protect our children, why not? One of my acquaintances received the 2nd dose. She is a girl of 14 years old. Everything is fine. Everybody there voluntarily vaccinates children. We always lag behind. They tell us to take a step ahead, but we take two steps back. Sadly. Therefore, we do not grow, and we do not develop
				negative	You must vaccinate yours!!! If you do not have brains, your children do not have one either!
2	26–07–2021	positive	We have 84% of our intensive care beds filled. They are loaded with patients who have not received vaccination against coronavirus infection and are now in a severe condition – 248 patients. Of these, 77 people are in extremely serious condition. This number scares us as doctors. We are reaching the peak that was last summer," said the head of the public health department of the capital, Timur Muratov.	negative	It is for those anti-vaccinators who can read and hear not only their cries about freedom. As soon as each of them understands the inhuman basis of personal freedom, opposed to the freedom of others, or rather other people, he / she is obliged to think.
				negative	The relatives of the deceased can sue the Shymkent anti-vaccinator (I forgot her name, sorry), which actively urges everyone to refuse vaccination
				negative	Let us gather money for the monuments to the killer doctors! Who sold out for premiums and killed people with the vaccine!!! They also lie!!! I am waiting for the heavenly punishment for you !!!
3	10–08–2021	positive	The first lot of the Chinese vaccine Sinopharm arrived in Kazakhstan on August 10, 2021. Following the negotiations with the People's Republic of China, an aircraft with the first batch of Sinopharm vaccine against coronavirus arrived in Almaty at the warehouses of the SK-Pharmacy Single Distributor.	positive	Good vaccine! It was recognized by WHO and Europe. Vaccinate. Health to all
				positive	Hayat is a good vaccine, so this one too. It is judging by my own example.
				negative	I doubt very much that WHO is responsible for our health and life.
				negative	Perhaps the quality of this vaccine is good (I do not argue). Just answer what it is made of, what is included in the composition?

				positive	Hooray, I'm going to put it. Do not miss the vaccine that came at the expense of the people.
4	11–08–2021	negative	The Ministry of Health of the Republic of Kazakhstan notes that 99.9% of the incidence of Covid-19 falls on unvaccinated citizens. In assessing the effectiveness of vaccination, it was found that 99.9% of the incidence of coronavirus infection falls on unvaccinated, while the proportion of patients after vaccination was only 0.1%, such data reported today by the Minister of Health Alexei Tsoi at a meeting of the Government.	negative	And it is true! Three friends are now in the hospital. There are no vaccinated people in the wards.
				negative	What is the percentage of re-illnesses? If such statistics do not even exist, then this means that there are no more patients, and then the question arises, why vaccinate those who have already been ill?
5	02–07–2021	negative	The "Indian" strain was found in all regions of Kazakhstan and the cities of Nur-Sultan, Almaty, Shymkent, zakon.kz reports. According to the Ministry of Health, the department carried out PCR screening of positive laboratory samples obtained from patients with coronavirus infection (CVI).	negative	There is no Indian strain. They said officially. It is ours who are lying to make people run to shoot up drugs. The day before yesterday, it was in 4 regions, and yesterday it was in all. Walked in the wind
				positive	Well, there is no point in getting vaccinated!
				negative	Do not write Indian. People in India know how upset it is
				negative	The Hindus themselves say there is no such thing.
				positive	These viruses appear abroad, but they come to us to die.

The experimental results have been extensively studied and analyzed to understand the root of the public's negative sentiment. Based on the data obtained by the OMSystem, it was concluded that Kazakhstanis, for the most part, do not trust the governmental methods of combating the pandemic. It should also be noted that users of social networks cannot identify fake news, trusting unverified information. Therefore, the experiment conducted on the topic of vaccination against the coronavirus disease makes it possible to understand the public's attitude and the Government's activities by assessing comments' sentiment analysis and semantic content. As a result, it will make it possible to maintain an exploratory policy for the public correctly, determine the presentation style of information material, accelerate the introduction of such large-scale state tasks, and ensure the preservation of public health. Furthermore, the OMSystem is used as a serious analytics tool to estimate the user perception of the social and economic life, which will allow quick explanations for the public, identify alarming factors of the public, and evaluate social mood.

7. Conclusion

A comparative analysis of foreign analytics platforms and the developed Kazakhstani OMSystem made it possible to conclude that foreign analytics platforms are mostly aimed at business and brand promotion. At the same time, they cover only the information space of foreign countries and are little focused on existing social problems. The analytics platforms of Kazakhstan, on the contrary, pay great attention to the analysis of public opinion on a wide range of political and socio-economic problems. They aim to cover the most relevant topics over large and small-time ranges and use ML algorithms to quickly and efficiently determine the sentiment of texts and user comments. The OMSystem monitors the current political and

socio-economic situation in the country, allows searching for the keywords on any desired topics, defines topics' sentiment with the dictionary and ML algorithms approaches, determines the social well-being based on such indicators as the level of topic discussion activity in society, the level of interest in the topic in society, and the level of social mood. The OMSystem's functionalities have been thoroughly demonstrated on the analysis of the vaccination against the coronavirus disease topic, providing the complex analytics in convenient graphics and tables.

Abbreviations

SA: sentiment analysis

ML: machine learning

SVM: support vector machine

NB: naïve Bayes

k-NN: k-nearest neighbors

LR: logistic regression

DT: decision tree

FR: random forest

CNN: convolutional neural networks

RNN: recurrent neural networks

LSTM: long short-term memory

REST API: Representational state transfer application programming interface

APIs: application programming interfaces

SMMM: social media marketing management

Declarations

Acknowledgements

We would like to thank the Science Fund for the commercialization of scientific and technical activities for providing the Grant №0101-18-GC on the topic "The opinion monitoring information system OMSystem."

Funding

This research was financed by the Grant of the Science Fund for the commercialization of scientific and technical activities №0101-18-GC on the topic "The opinion monitoring information system OMSystem."

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The database of the experimental results of the OMSystem, corresponding pivot tables in the Russian language, and their translation to the English language and graphics are available from the corresponding author on the request.

Competing interests

The authors declare that they have no competing interests.

Contributions

V.K. designed the content of the research paper, described the obtained experimental results, and wrote the main manuscript text. G.M. and Z.M. designed the pivot tables from the experimental data and checked the paper's writing. G.N. analyzed social-economic and political aspects of received experimental results and made conclusions based on the public's opinion. S.T. and Z.S. configured the OMSystem and ran the experiment. Finally, M.N. made a thorough revision of the paper, corrected it, and proposed the ways of its significant improvement. All authors reviewed the manuscript.

References

1. Zhang H, Zang Zh, Zhu H, Uddin MI, Amin MA. Big data-assisted social media analytics for business model for business decision making system competitive analysis. Information Processing & Management. 2022; 59(1): 1–12. https://doi.org/10.1016/j.ipm.2021.102762.
2. Singh H, Yadav A, Bansal R, Mala S. Understanding Brand Authenticity Sentiments using Big Data Analytics, in 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence), 2021, pp. 304–308, https://doi.org/10.1109/Confluence51648.2021.9377046.
3. Pamuksuz U, Yun JT, Humphreys A. A Brand-New Look at You: Predicting Brand Personality in Social Media Networks with Machine Learning. Journal of Interactive Marketing. 2021; 56: 55–69. https://doi.org/10.1016/j.intmar.2021.05.001.
4. Praet S, Van Aelst P, van Erkel P. Van der Veeken S, Martens D. Predictive modeling to study lifestyle politics with Facebook likes. EPJ Data Science. 2021; 10 (50). https://doi.org/10.1140/epjds/s13688-021-00305-7.
5. Chandra Sekhar Reddy N, Subhashini V, Rai D, Sriharsha, Vittal B, Ganesh S. Product rating estimation using machine learning, in 6th International Conference on Communication and Electronics Systems (ICCES), 2021, pp. 1366–1369, https://doi.org/10.1109/ICCES51350.2021.9489208.
6. Dangi D, Bhagat A, Dixit DK. Emerging Applications of Artificial Intelligence, Machine learning and Data Science. CMC-Computers, Materials & Continua. 2022; 70(3): 5399–5419. https://www.techscience.com/cmc/v70n3/44963.
7. Qureshi MA, Asif M, Hassan MF, Mustafa G, Ehsan MK, Ali A, Sajid U. A Novel Auto-Annotation Technique for Aspect Level Sentiment Analysis. CMC-Computers, Materials & Continua. 2022; 70(3): 4987–5004. https://www.techscience.com/cmc/v70n3/44970.
8. Benedetto F, Tedeschi A. Big Data Sentiment Analysis for Brand Monitoring in Social Media Streams by Cloud Computing. In: Pedrycz W, Chen SM, editors. Sentiment Analysis and Ontology Engineering. Studies in Computational Intelligence. Springer: Cham; 2016. pp. 341–377. https://doi.org/10.1007/978-3-319-30319-2_14.
9. Schinas M, Papadopoulos S, Apostolidis L, Kompatsiaris Y, Mitkas PA. Open-Source Monitoring, Search and Analytics Over Social Media. In: Kompatsiaris I et al., editors. Internet Science. INSCI 2017. Lecture Notes in Computer Science. Springer: Cham; 2017. pp. 361–369. https://doi.org/10.1007/978-3-319-70284-1_28.
10. Radicioni T, Saracco F, Pavan E., Squartini T. Analysing Twitter semantic networks: the case of 2018 Italian elections. Scientific Reports. 2021; 11: 13207. https://doi.org/10.1038/s41598-021-92337-2.
11. Bhatnagar S, Choubey N. Making sense of tweets using sentiment analysis on closely related topics. Social Network Analysis and Mining. 2021; 11(44). https://doi.org/10.1007/s13278-021-00752-0.
12. Nandwani P, Verma R. A review on sentiment analysis and emotion detection from text. Social Network Analysis and Mining. 2021; 11(81). https://doi.org/10.1007/s13278-021-00776-6.

13. Hartmann J, Huppertz J, Schamp C, Heitmann M. Comparing automated text classification methods. International Journal of Research in Marketing. 2021; 36(1): 20–38. https://doi.org/10.1016/j.ijresmar.2018.09.009.
14. Huq MR, Ali A, Rahman A. Sentiment Analysis on Twitter Data using KNN and SVM. International Journal of Advanced Computer Science and Applications (IJACSA). 2017; 8(6). http://dx.doi.org/10.14569/IJACSA.2017.080603.
15. Dang NC, Moreno-García MN, De la Prieta F. Sentiment Analysis Based on Deep Learning: A Comparative Study. Electronics. 2020; 9(3): 483. https://doi.org/10.3390/electronics9030483.
16. Röchert D, Neubaum G, Stieglitz S. Identifying Political Sentiments on YouTube: A Systematic Comparison Regarding the Accuracy of Recurrent Neural Network and Machine Learning Models. In: van Duijn M, Preuss M, Spaiser V, Takes F, Verberne S, editors. Disinformation in Open Online Media. MISDOOM 2020. Lecture Notes in Computer Science. Springer: Cham; 2020. pp. 107-121. http://dx.doi.org/10.1007/978-3-030-61841-4_8.
17. Ombabi AH, Ouarda W, Alimi AM. Deep learning CNN–LSTM framework for Arabic sentiment analysis using textual information shared in social networks. Social Network Analysis and Mining. 2020; 10(53). https://doi.org/10.1007/s13278-020-00668-1.
18. Znovarev A, Bilyi A. A comparison of machine learning methods of sentiment analysis based on Russian language Twitter data. In: 11th Majorov International Conference on Software Engineering and Computer Systems, MICSECS, Saint Petersburg, Russian Federation. 2020, pp. 1–7.
19. Hamada MA, Sultanbek K, Alzhanov B, Tokbanov B. Sentimental text processing tool for Russian language based on machine learning algorithms. In: Proceedings of the 5th International Conference on Engineering and MIS, Astana, Kazakhstan. 2019, pp. 1–6.
20. Yergesh B, Bekmanova G, Sharipbay A. Sentiment analysis of Kazakh text and their polarity. Web Intelligence. 2019; 17(1): 9–15.
21. Bekmanova G, Yelibayeva G, Aubakirova S, Dyussupova N, Sharipbay A, Niyazova N. Methods for analyzing polarity of the Kazakh texts related to the terrorist threats. In: 19th International Conference on Computational Science and Its Applications, ICCSA, Saint Petersburg, Russian Federation. 2019, pp. 717–730.
22. Alzahrani H, Acharya S, Duverger P, Nguyen Nam P. Contextual polarity and influence mining in online social networks. Computational Social Networks. 2021; 8(21). https://doi.org/10.1186/s40649-021-00101-3.
23. Weber D, Nasim M, Mitchell L, Falzon L. Exploring the effect of streamed social media data variations on social network analysis. Social Network Analysis and Mining. 2021; 11(62). https://doi.org/10.1007/s13278-021-00770-y.
24. Sproutsocial. https://sproutsocial.com/. Accessed 27 November 2021.
25. Chaudhary K, Alam M, Al-Rakhami MS, Gumaei A. Machine learning-based mathematical modelling for prediction of social media consumer behavior using big data analytics. Journal of Big Data. 2021; 8(73). https://doi.org/10.1186/s40537-021-00466-2.
26. Hubspot. https://www.hubspot.com/. Accessed 27 November. 2021.
27. Buzzsumo. https://buzzsumo.com/. Accessed 27 November. 2021.
28. Deng Q, Hine MJ, Ji Sh, Wang Y. Understanding consumer engagement with brand posts on social media: The effects of post linguistic styles. Electronic Commerce Research and Applications. 2021; 48. https://doi.org/10.1016/j.elerap.2021.101068.
29. Hootsuite. https://www.hootsuite.com/. Accessed 27 November 2021.
30. Brandmention. https://brandmentions.com/. Accessed 27 November 2021.
31. Rahmatulloh A, Shofa RN, Darmawan I, Ardiansah, Sentiment Analysis of Ojek Online User Satisfaction Based on the Naïve Bayes and Net Brand Reputation Method. In: 9th International Conference on Information and Communication Technology (ICoICT). 2021, pp. 337–341. https://doi.org/10.1109/ICoICT52021.2021.9527466.
32. IQBuzz. https://iqbuzz.pro/. Accessed 27 November 2021.
33. Beltrán J, Jara-Reyes R, Faure A. The Emotions of the Outbreak. Topics, Sentiments and Politics on Twitter During Chilean October. In: Rocha Á, Barredo D, López-López PC, Puentes-Rivera I, editors. Communication and Smart Technologies.

ICOMTA 2021. Smart Innovation, Systems and Technologies. Springer: Singapore; 2022, pp. 216–226. https://doi.org/10.1007/978-981-16-5792-4_22.

34. Snaplytics. https://thehub.io/startups/snaplytics. Accessed 27 November 2021.
35. iMAS. https://imas.kz/. Accessed 27 November 2021.
36. Alem Media Monitoring. https://alem.kz/en/monitoring-smi/. Accessed 27 November 2021.
37. Usero B, Hernández V, Quintana C. Social Media Mining for Business Intelligence Analytics: An Application for Movie Box Office Forecasting. In: Arai K, editors. Intelligent Computing. Lecture Notes in Networks and Systems. Springer: Cham; 2022, pp. 981–999. https://doi.org/10.1007/978-3-030-80119-9_65.
38. Mutanov G, Karyukin V, Mamykova Zh. Multi-Class Sentiment Analysis of Social Media Data with Machine Learning Algorithms. CMC–Computers, Materials & Continua. 2021; 69(1): 913–930. https://www.techscience.com/cmc/v69n1/42767.
39. Kadyrbek N, Sundetova Zh, Torekul S. Information Monitoring System of Social Wellness Opinions. In: IEEE 8th Workshop on Advances in Information, Electronic and Electrical Engineering (AIEEE). 2021, pp. 1-4. https://doi.org/10.1109/AIEEE51419.2021.9435811.
40. Mussiraliyeva Sh, Omarov B, Yoo P, Bolatbek M. Applying Machine Learning Techniques for Religious Extremism Detection on Online User Contents. CMC-Computers, Materials & Continua. 2021; 70(1): 915–934. https://www.techscience.com/cmc/v70n1/44387.
41. Khalil EAH, Houby EMFE & Mohamed HK. Deep learning for emotion analysis in Arabic tweets. Journal of Big Data. 2021; 8(136). https://doi.org/10.1186/s40537-021-00523-w.
42. Domalewska D. An analysis of COVID-19 economic measures and attitudes: evidence from social media mining. Journal of Big Data. 2021; 8(42). https://doi.org/10.1186/s40537-021-00431-z.
43. Bhatnagar S, Choubey N. Making sense of tweets using sentiment analysis on closely related topics. Social Network Analysis and Mining. 2021; 11(44). https://doi.org/10.1007/s13278-021-00752-0.
44. Ramya GR, Bagavathi Sivakumar P. An incremental learning temporal influence model for identifying topical influencers on Twitter dataset. Social Network Analysis and Mining. 2021; 11(27). https://doi.org/10.1007/s13278-021-00732-4.
45. Heidari M, Shamsinejad P. Producing An Instagram Dataset For Persian Language Sentiment Analysis Using Crowdsourcing Method. In: 6th International Conference on Web Research (ICWR). 2020, pp. 284–287. https://doi.org/10.1109/ICWR49608.2020.9122270.
46. Camacho D, Luzón MV, Cambria E. New research methods & algorithms in social network analysis. Future Generation Computer Systems. 2021; 114: 290–293. https://doi.org/10.1016/j.future.2020.08.006.
47. Zarzour H, Al shboul B, Al-Ayyoub M, Jararweh Y. Sentiment Analysis Based on Deep Learning Methods for Explainable Recommendations with Reviews. In: 12th International Conference on Information and Communication Systems (ICICS). 2021, pp. 452–456. https://doi.org/10.1109/ICICS52457.2021.9464601.
48. Ng LHX, Loke JY. Analyzing Public Opinion and Misinformation in a COVID-19 Telegram Group Chat. *IEEE Internet Computing*. 2021; 25(2): 84–91. doi: https://doi.ieeecomputersociety.org/10.1109/MIC.2020.3040516.
49. Kulchitskaya DY, Folts AO. Between politics and show business: Public discourse on social media regarding ksenia sobchak, the only female candidate in the 2018 Russian presidential election. Monitoring Obshchestvennogo Mneniya: Ekonomicheskie i Sotsial'nye Peremeny. 2020; 4:176–199 https://doi.org/10.14515/monitoring.2020.4.1254.
50. Engagement Rate: A Metric You Can Count On. https://www.socialbakers.com/blog/1427-engagement-rate-a-metric-you-can-count-on. Accessed 27 November 2021.
51. Porreca A, Scozzari F, Di Nicola M. Using text mining and sentiment analysis to analyze YouTube Italian videos concerning vaccination. BMC Public Health. 2020; 259. https://doi.org/10.1186/s12889-020-8342-4.
52. Karami A, Zhu M, Goldschmidt B, Boyajieff HR, Najafabadi MM. COVID-19 Vaccine and Social Media in the US: Exploring Emotions and Discussions on Twitter. Vaccines. 2021; 9(10): 1059. https://doi.org/10.3390/vaccines9101059.

53. Nasir A, Ali Shah M, Ashraf U, Khan A, Jeon G. An intelligent framework to predict socio-economic impacts of COVID-19 and public sentiments. Computers & Electrical Engineering. 2021; 96. https://doi.org/10.1016/j.compeleceng.2021.107526.
54. Singh M, Jakhar AK, Pandey S. Sentiment analysis on the impact of coronavirus in social life using the BERT model. Social Network Analysis and Mining. 2021; 11(33). https://doi.org/10.1007/s13278-021-00737-z.
55. Marcec R, Likic R. Using Twitter for sentiment analysis towards AstraZeneca/Oxford, Pfizer/BioNTech and Moderna COVID-19 vaccines. Postgraduate Medical Journal. 2021. doi: 10.1136/postgradmedj-2021-140685.
56. Sahraian MA, Ghadiri F, Azimi A, Moghadasi AN. Adverse events reported by Iranian patients with multiple sclerosis after the first dose of Sinopharm BBIBP-CorV. Vaccine. 2021; 39(43): 6347-6350. https://doi.org/10.1016/j.vaccine.2021.09.030.

Figures

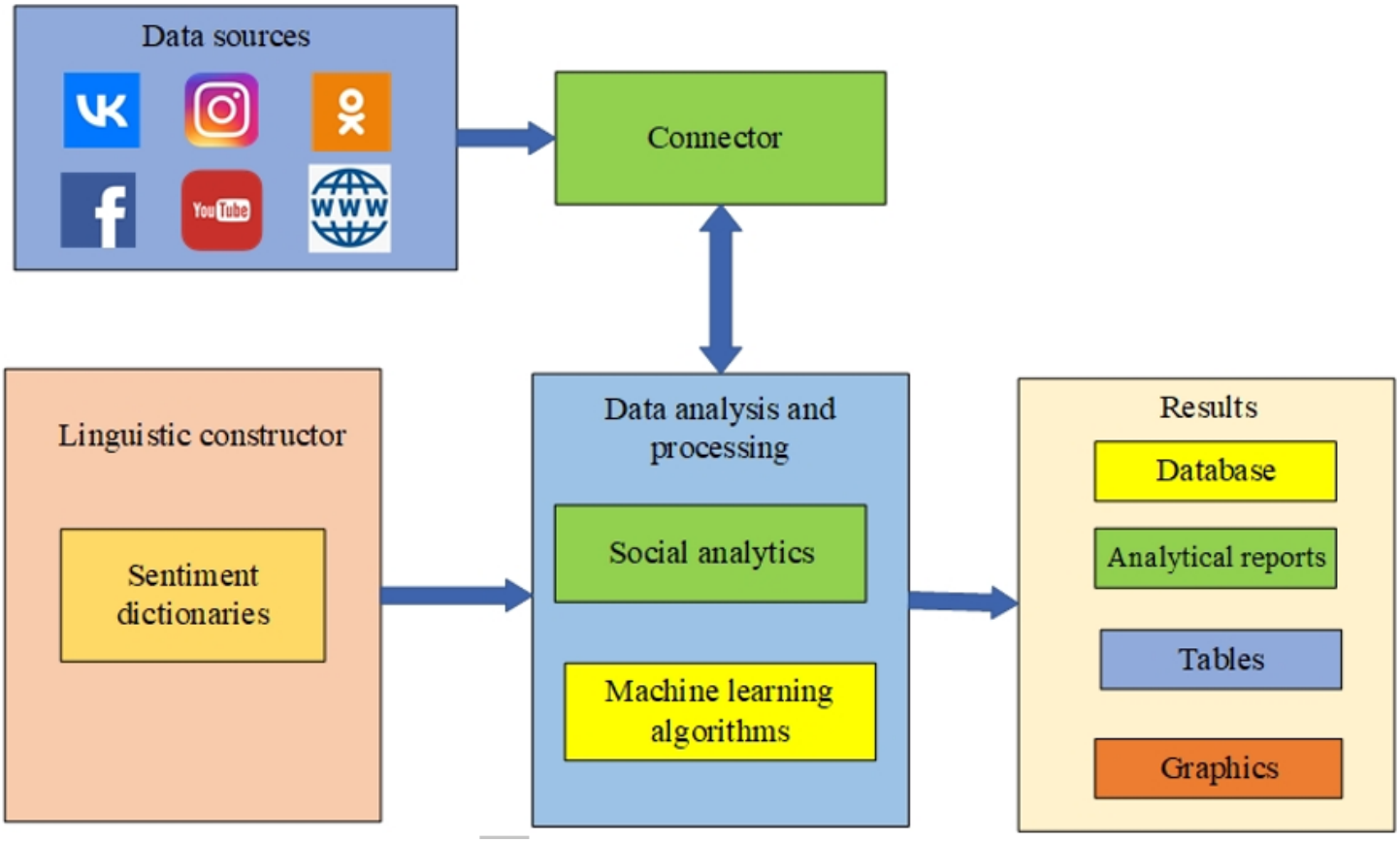

Figure 1

System Architecture

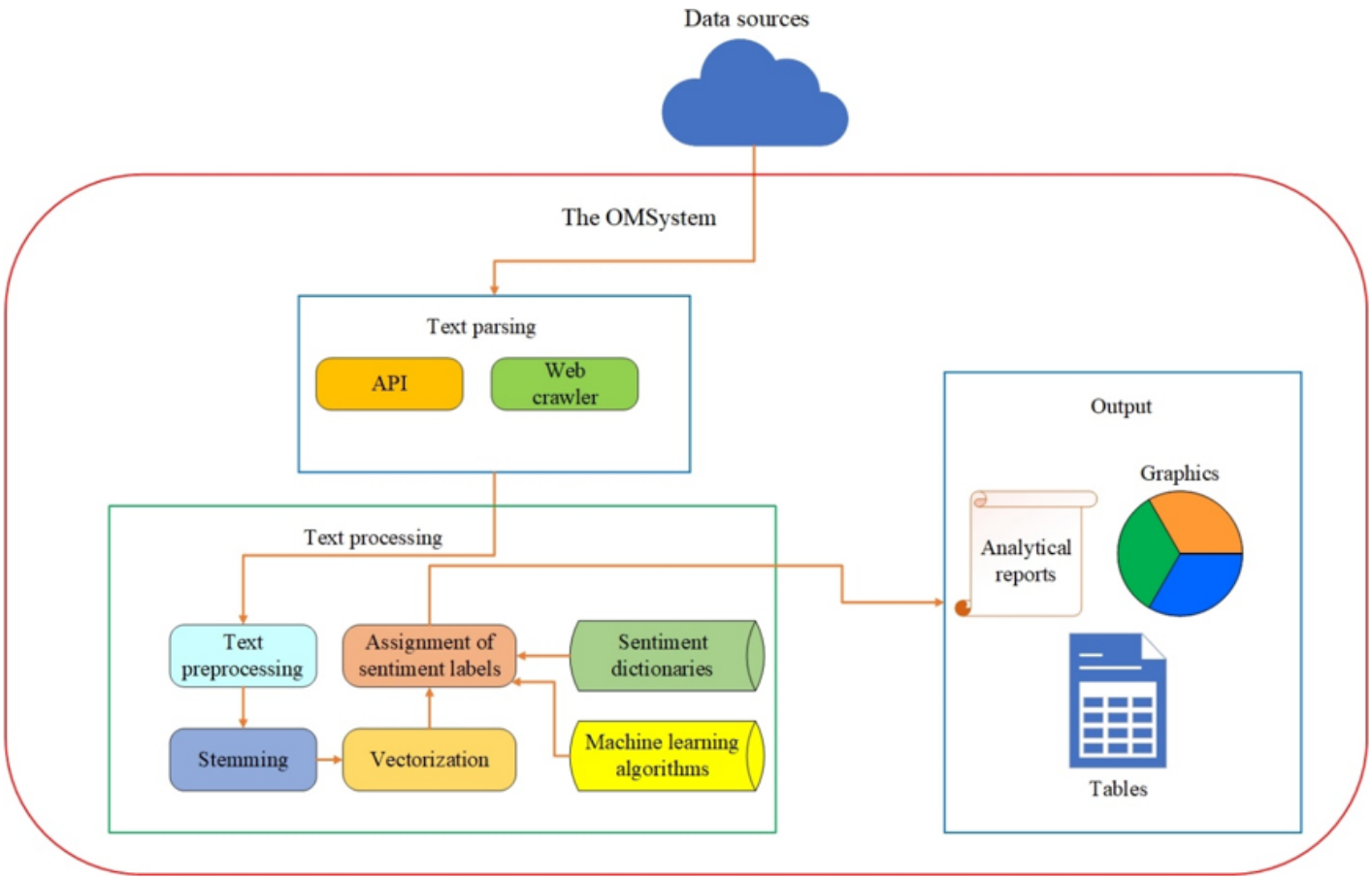

Figure 2

The OMSystem's analytics building steps

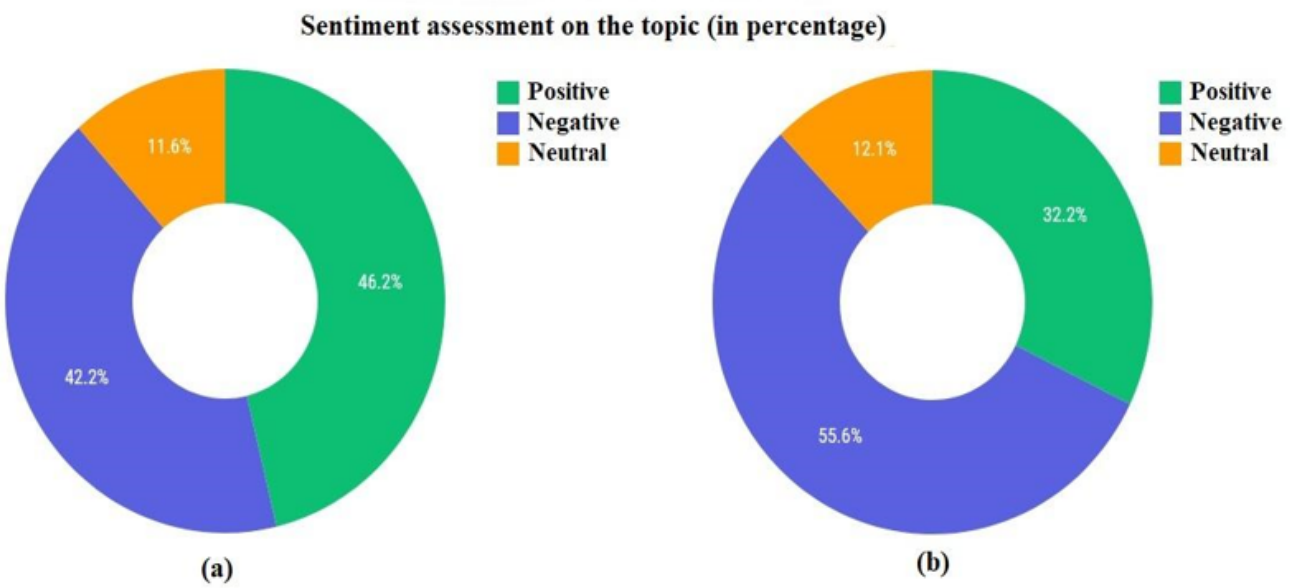

Figure 3

Evaluation of the sentiment of the first period – (a) Almaty and Nur-Sultan, (b) large regional cities

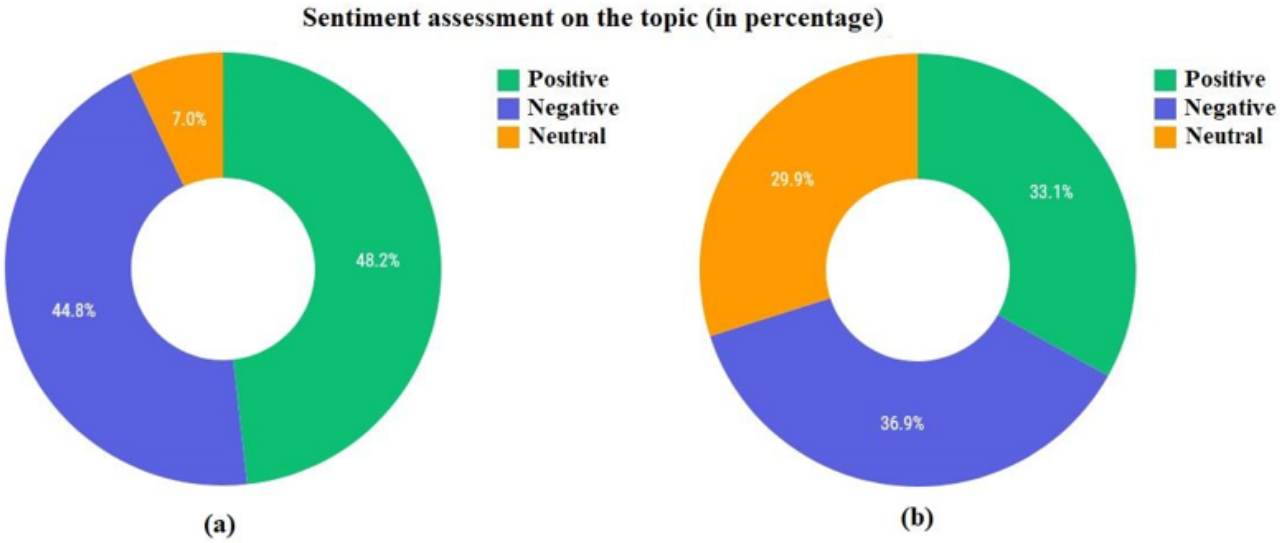

Figure 4

Evaluation of the sentiment of the second period – (a) Almaty and Nur-Sultan, (b) large regional cities