

ICEST 2022**III International Conference on Economic and Social Trends for Sustainability of Modern Society****MONITORING OF INFORMATION EVENTS FOR THE SOCIO-CULTURAL BACKGROUND ANALYSIS DURING EPIDEMICS**

Sergey Kuleshov (a), Alexandra Zaytseva (b)*, Alexey Aksenov(c)

*Corresponding author

(a) St. Petersburg Federal Research Center of the Russian Academy of Sciences, St. Petersburg, 199178, Russia,
kuleshov@ias.spb.su(b) St. Petersburg Federal Research Center of the Russian Academy of Sciences, St. Petersburg, 199178, Russia,
cher@ias.spb.su(c) St. Petersburg Federal Research Center of the Russian Academy of Sciences, St. Petersburg, 199178, Russia,
a_aksenov@ias.spb.su**Abstract**

The article is devoted to the study of the issues of forecasting the impact of epidemics and epidemiological informing measures on the psychological state of society by monitoring information events and analyzing the sociocultural background of the population. The aim of the work is to analyze and develop methods of content analysis for a system for monitoring the psychological state of society on information reactions on the Internet, including methods for sentiment analysis of user comments in social networks, using an associative-ontological approach to form a sample of analyzed comments in one thematic area. As a result of the study a monitoring system was developed, built on a microservice architecture and including an information event monitoring server, containing a data collection platform from the Internet, sentiment analysis linguistic services, an ElasticSearch database that stores processed texts in an indexed and searchable form, researcher's workplace with a web interface for setting up tasks and viewing the results of work. The system developed provides additional indicators to refine new models for substantiating epidemiological awareness programs for the population in large metropolitan areas connected by passenger flows of different intensity, as well as the ability to predict the impact of epidemics and epidemiological awareness activities on the psychological state of society by monitoring information events and analyzing the sociocultural background of the population.

2357-1330 © 2022 Published by European Publisher.

Keywords: Content analysis, epidemics, monitoring of information events, psychological state of society

1. Introduction

During periods of mass epidemics or pandemics the monitoring of the psychological state of society becomes of particular importance for a correct understanding of the essence of social phenomena and the construction of adequate models of the possible development of the epidemiological situation both in a single region and throughout the country (or the world). In the context of the global digitalization of society, especially in megacities, where crowding is one of the potential factors for the rapid spread of viruses, the best way to track the psychological state of society is to monitor social networks in terms of analyzing the reactions (and comments) of users to a particular information event, tracking profiles both individual users and various kinds of communities. Taking into account the data obtained from such monitoring and proactive response to a changing situation based on model forecasts makes it possible to level the consequences of epidemics for modern society. Another positive effect of such monitoring is the possibility of early detection of new sociocultural phenomena formed under the influence of various factors, for example, measures to counteract the spread of epidemics. There are a fairly large number of approaches to the analysis of the psychological state of society in the context of epidemics, for example, in (Anosov, 2021) the main milestones in the transformation of public attention, which forms the general psychological mood of social groups, are analyzed. In (Abdullin et al., 2021; Bessi et al., 2015; Deineka & Maksimenko, 2020; Depoux et al., 2020; Malysheva, 2020), the impact of the infodemic on the public moods of social media users is analyzed, including using content-based methods. analysis. An increased number of studies on the impact of the reactions of social network users on the general psychological state of society, for example (Ahmed et al., 2020; Arkhipova et al., 2020; Basch et al., 2020; Bazanov & Gubanov, 2013; Dementieva et al., 2020; Martyshkin & Pashchenko, 2019; Petrov et al., 2020; Vasilkova & Legostaeva, 2020) emphasizes the relevance of the problem described.

2. Problem Statement

The task of the study is to search for the patterns of social processes and their reflection in the form of a reaction to information events. The study consists in processing reactions by collecting user comments on social networks with their temporary localization, as well as quantitative indicators of reactions to them (number of views, likes, reposts), carried out during the active phase of the pandemic in Russia.

To solve the problem of developing a system for monitoring information events and analyzing the sociocultural background of the Russian-speaking population, it is proposed to use a combination of data-retrieval methods from Internet resources, sentiment analysis of Russian-language texts with methods of an associative-ontological approach to processing data in natural language.

3. Research Questions

The article is devoted to the study of the issues of predicting the impact of epidemics and epidemiological information measures on the psychological state of society by monitoring information events and analyzing the sociocultural background of the population.

It is planned to answer the following questions:

- if the emotional reaction of the population depends on the types of information events;
- if the incidence curve during a pandemic correlates with the level of emotional reactions of users;
- if the emotional state of users and, accordingly, the psychological state of society depend on the measures introduced.

4. Purpose of the Study

The purpose of the work is to analyze and develop content analysis methods for a system for monitoring the psychological state of society by information reactions on the Internet, including methods for sentiment analysis of user comments in social networks, using the associative-ontological approach when forming the analyzed sample (Kuleshov et al., 2015; Kuleshov et al., 2020).

5. Research Methods

To achieve the goal of the study an analysis was made of the existing methods of content analysis of Russian-language texts, their advantages and disadvantages.

The following methods and services, both commercial and research, were analyzed: Brand24, SentiStrength, the sentiment analysis component of the Analytical Courier and X-files systems, Vaal, the sentiment analysis component of the RCO Fact Extractor system, Linis Crowd, Tweenator, Sentiment Analysis of Tweets using Multinomial Naive Bayes.

The Brand24 software (<https://www.syssoft.ru/Brand24/>) is a proprietary commercial product that enables social media monitoring, Internet monitoring, marketing and social media analysis. that allows you to identify and analyze in real time online conversations about brands or products. For the purposes of the study, it is unsuitable due to the impossibility of access to the source program code, as well as the limitation of the volume of processed data. System developed by Thelwall, Buckley, Paltoglou and Cai. SentiStrength (<http://sentistrength.wlv.ac.uk/>) is designed to analyze short unstructured informal texts in English, but it can be configured to work with texts in other languages, including Russian. The result is given as an assessment of the positive and negative components on a scale from 1 to 5 with the corresponding sign. The algorithm is based on finding the maximum sentiment value in the text for each scale (i.e. searching for the word with the maximum negative score and the word with the maximum positive score). The algorithm takes into account the simplest interaction of words and idiomatic expressions (Novikov et al., 2018). The use of the system for the analysis of Russian-language texts is complicated by the fact that its algorithms do not take into account the specifics of the morphology of the Russian language, and sentiment is assessed for the entire text as a whole, without highlighting subjects

and objects of sentiment. The sentiment analysis component of the Analytical Courier and X-files systems implements a method based on dictionaries and rules. (https://www.tadviser.ru/index.php/Product:Analytical_courier). Sentiment is assessed on a ternary scale (positive/negative/neutral). The system does not quantify sentiment. The Vaal system (<http://www.vaal.ru/>) is designed to assess the "unconscious emotional impact of the phonetic structure of the text and individual words on the human subconscious." The operation of the system is based on turning the text into a frequency dictionary and assigning some words to certain psycholinguistic categories. The result of the analysis is given to the user in the form of a set of ratings for a number of criteria related to the given text/word. The system does not analyze the semantics of the text, which leads to a very limited applicability of the product, the use of the system in automatic mode is impossible. The sentiment analysis component of the RCO Fact Extractor system (<http://www.rco.ru/>) uses a rule-based approach, takes into account the syntactic structure of the text and the interaction of different types of words. There is no quantitative assessment of the tone of the text.

The Sentiment Analysis of Tweets in Russian using Multinomial Naive Bayes (Smetanin, 2020) method was chosen of all the analyzed methods for the implementation of content analysis in the system under development. It provides the most optimal results for assessing the sentiment of short Russian-language messages; out of 17,639,674 records (Rubtsova, 2012).

The developed monitoring system is built on a microservice architecture (Figure 1) and includes an information event monitoring server containing a platform for collecting data from the Internet (Kuleshov et al., 2020), linguistic services, including sentiment analysis; a database based on ElasticSearch, which stores the processed texts in an indexed and searchable form; researcher's workplace with a web interface for setting up tasks and viewing the results of work.

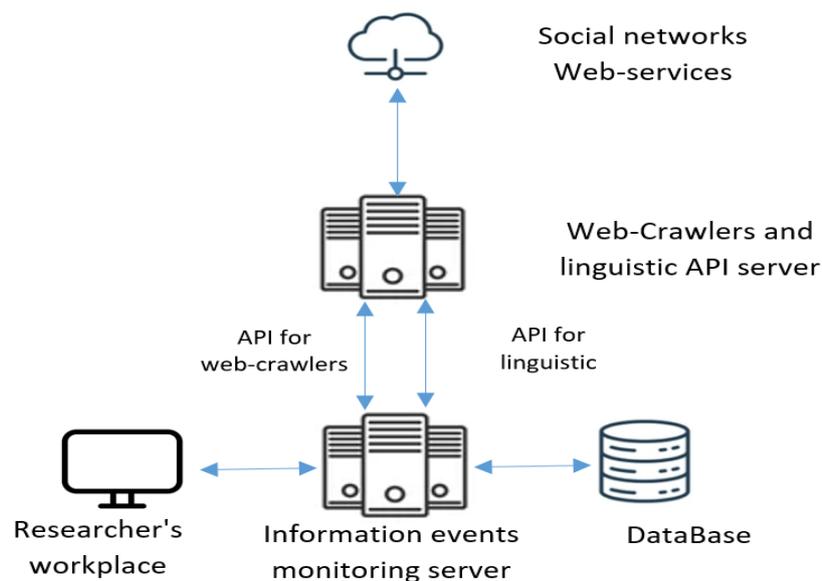


Figure 1. The structure of the system for monitoring information events based on the analysis of thematic channels in social networks

6. Findings

As a result of the operation of the information events monitoring system based on the analysis of thematic channels in social networks a corpus of comments from social network users on news events in the official accounts @rospotrebnadzor.official, @stopcoronavirusrf for the period from March 2020 to April 2022 was obtained.

The essence of the experiment is to collect messages (descriptions of information events), user comments to them, as well as quantitative indicators of reactions to them (the number of views, likes, reposts), carried out during the active phase of the pandemic in Russia from May 2020 to April 2022. Message texts undergo technical and linguistic pre-processing and are stored in a database; for each message and comment, sentiment analysis is performed (Smetanin, 2020) to obtain summary numerical characteristics of the reaction to information events that day and determine the socio-cultural background of the population.

Official channels on social networks @rospotrebnadzor.official, @stopcoronavirusrf and a number of others serve as sources of information events and comments.

Figures 2-4 show the results of the monitoring system.

Figure 2 shows the ratio of positive and negative comments on information events on the topic of coronavirus in a given period of time. Figure 3 shows the distribution of "likes" by information events.

Figure 4 is provided to compare the incidence statistics with the given reactions of users to events.

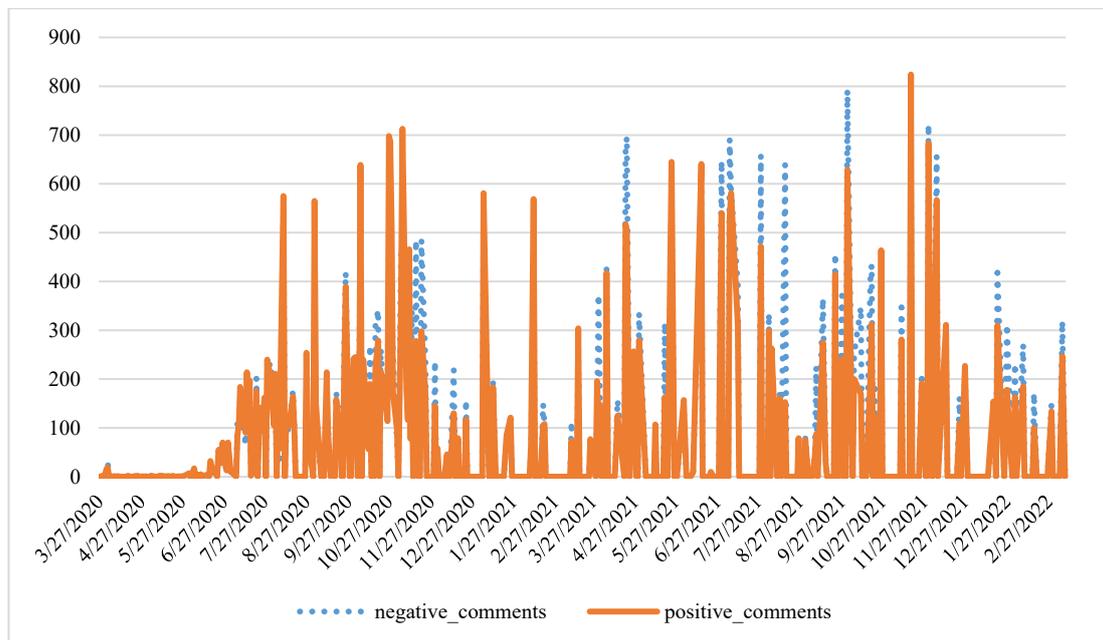


Figure 2. Ratio of positive and negative comments on news events related to coronavirus

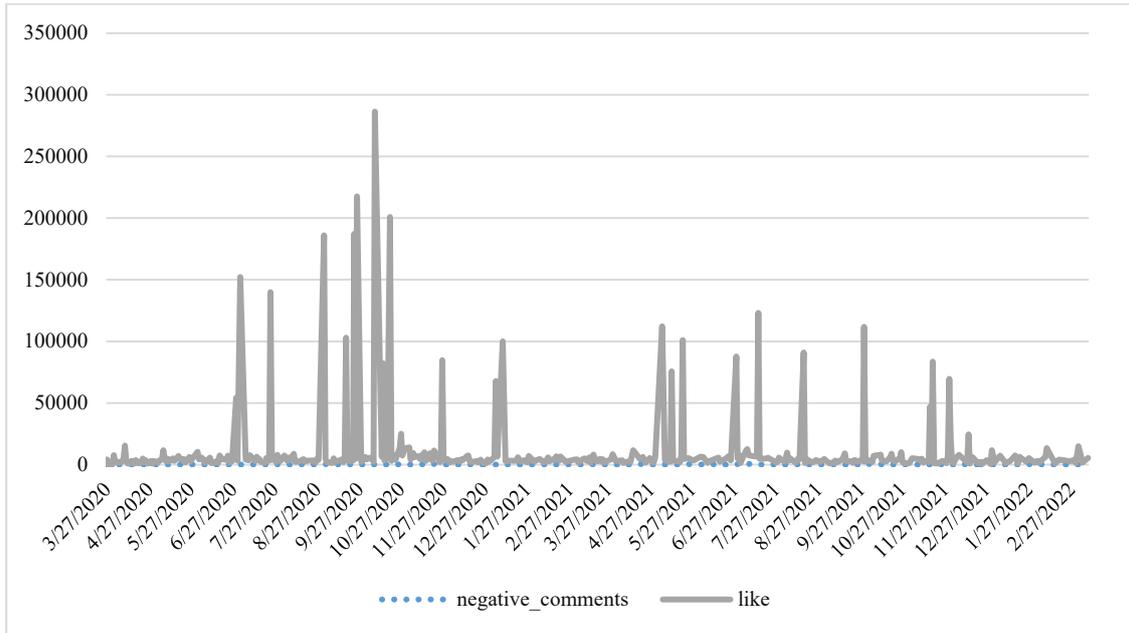


Figure 3. Distribution of "likes" for informational events against the background of negative comments

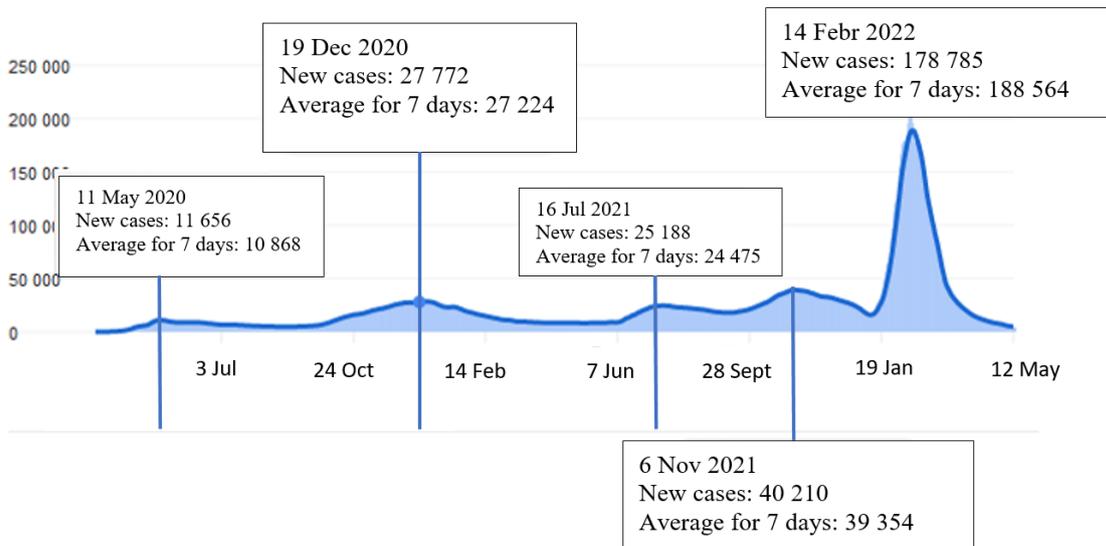


Figure 4. Statistics of new cases of coronavirus in Russia for the entire period of the epidemic (JHU CSSE COVID-19 Data, 2022)

When analyzing Figures 2-4 we can say that outbursts of social reaction (both negative and positive) occur approximately in the same periods as the outbreak of morbidity, that is, they reflect the natural reaction of the population to information events in the space of social networks, most sharp bursts of negative emotions correspond to days in which there was awareness of the need to comply with sanitary standards, tightening restrictions on mass events and the use of vaccines.

7. Conclusion

As a result of the study content analysis methods were analyzed and developed for the system for monitoring the psychological state of society by information reactions on the Internet, including methods for sentiment analysis of user comments in social networks, using the associative-ontological approach when forming the analyzed sample. The structure of the system for monitoring information events based on the analysis of thematic channels in social networks has been developed.

The experiments carried out as a result of the study using the developed monitoring system made it possible to obtain answers to the questions posed. It was revealed that the emotional reactions of the population and the morbidity curve during the development of epidemics are quite strongly correlated, the strongest emotional reaction to information events associated with the measures and restrictions being introduced was found.

The system developed provides additional indicators to refine the method and models proposed in (Levashkin et al., 2021; Osipov et al., 2022) for substantiating epidemiological awareness programs for the population in large metropolitan areas connected by passenger flows of different intensity, as well as the possibility of predicting the impact of epidemics and epidemiological awareness activities on the psychological state of society by monitoring information events and analyzing the socio-cultural background of the population.

Acknowledgments

This study was funded by the Russian Foundation for Basic Research, project number 20-04-60455.

References

- Abdullin, A. G., Likholeto, V. V., & Rylskaya, E. A. (2021). Razmyshleniya o pol'ze pandemii dlya moral'no-psihologicheskogo ozdorovleniya obshchestva [Reflections on the benefits of a pandemic for the moral and psychological recovery of society]. *Psychopedagogy in law enforcement*, 26(2) (85), 144-155. <https://doi.org/10.24412/1999-6241-2021-2-144-155>
- Ahmed, M. Z., Ahmed, O., Aibao, Z., Hanbin, S., Siyu, L., & Ahmad, A. (2020). Epidemic of COVID-19 in China and associated Psychological Problems. *Asian Journal of Psychiatry*, 51, 102092. <https://doi.org/10.1016/j.ajp.2020.102092>
- Anosov, S. S. (2021). Social'no-psihologicheskoe sostoyanie rossijskogo obshchestva [Socio-psychological state of the Russian society]. *Sociology*, 3, 5-12. <http://soziologi.ru/upload/iblock/219/7c71v6knk6hkuy66870robzqc1jfyu62/%E2%84%963%202021.pdf>
- Arkhipova, A. S., Radchenko, D. A., Kozlova, I. V., Peigin, B. S., Gavrilova, M. V., & Petrov, N. V. (2020). Puti rossijskoj infodemii: ot WhatsApp do Sledstvennogo komiteta [Ways of the Russian infodemic: from Whatsapp to the investigative committee]. *Monitoring Public Opinion: Economic and Social Change*, 6(160), 231-265. <https://doi.org/10.14515/monitoring.2020.6.1778>
- Basch, C. H., Hillyer, G. C., Meleo-Erwin, Z. C., Jaime, C., Mohlman, J., & Basch, C. E. (2020). Preventive Behaviors Conveyed on YouTube to Mitigate Transmission of COVID-19: Cross-Sectional Study. *JMIR Public Health and Surveillance*, 6(2), e18807. <https://doi.org/10.2196/18807>
- Bazenkov, N. I., & Gubanov, D. A. (2013). Obzor informacionnyh sistem analiza social'nyh setej [Review of information systems for social network analysis]. *Management of large systems: a*

- collection of papers, 41, 357-394.
http://ubs.mtas.ru/archive/search_results_new.php?publication_id=18967
- Bessi, A., Coletto, M., Davidescu, G. A., Scala, A., Caldarelli, G., & Quattrociocchi, W. (2015). Science vs Conspiracy: Collective Narratives in the Age of Misinformation. *PLOS ONE*, 10(2), e0118093. <https://doi.org/10.1371/journal.pone.0118093>
- Deineka, O. S., & Maksimenko, A. A. (2020). Ocenka psihologicheskogo sostoyaniya obshchestva v usloviyah infodemii na osnove analiza social'nyh setej: obzor zarubezhnyh publikacij [Assessing the psychological state of society in the context of an infodemic through the analysis of social networks: a review of foreign publications]. *Society. Wednesday. Development (Terra Humana)*, 2(55), 28-39.
- Dementieva, K. V. (2020). Mediakommunikacii regiona v usloviyah rasprostraneniya koronavirusa: osobennosti medijnoj povestki i вовлечeniya auditorii [Media communications of the region in the context of the spread of coronavirus: features of the media agenda and audience involvement]. *Humanitarian vector*, 15(5), 166-175. <https://doi.org/10.21209/1996-7853-2020-15-5-166-175>
- Depoux, A., Martin, S., Karafillakis, E., Preet, R., Wilder-Smith, A., & Larson, H. (2020). The pandemic of social media panic travels faster than the COVID-19 outbreak. *Journal of Travel Medicine*, 27(3). <https://doi.org/10.1093/jtm/taaa031>
- JHU CSSE COVID-19 Data. (2022). Retrieved on 12 May, 2022 from <https://coronavirus.jhu.edu/map.html>
- Kuleshov, S. V., Zaitseva, A. A., & Levashkin, S. P. (2020). Tekhnologii i principy sbora i obrabotki nestrukturirovannyh raspredelennyh dannyh s uchetom sovremennyh osobennostej predstavleniya media-kontenta [Technologies and principles for the collection and processing of unstructured distributed data, taking into account modern features of the provision of media content]. *Informatization and communication*, 4, 62-66. <https://doi.org/10.34219/2078-8320-2020-11-5-22-28>
- Kuleshov, S. V., Zaitseva, A. A., & Markov, S. V. (2015). Associativno-ontologicheskij podhod k obrabotke tekstov na estestvennom yazyke [Associative-ontological approach to natural language text processing]. *Intelligent technologies in transport*, 4, 40-45. <http://archive.it-pgups.ru/issues/2015/04/ITT-2015-04-associative-ontological-approach.pdf>
- Levashkin, S. P., Zakharova, O. I., Kuleshov, S. V., & Zaytseva, A. A. (2021). Adaptive-compartmental model of coronavirus epidemic and its optimization by the methods of artificial intelligence. *Journal of Physics. Conference Series*, 1864, 012108. <https://doi.org/10.1088/1742-6596/1864/1/012108>
- Malysheva, G. A. (2020). Social'no-politicheskie aspekty pandemii v obshchestve cifrovyyh setej: opyt Rossii [Socio-political aspects of the pandemic in the society of digital networkization: Russian experience]. *Vestnik of MGOU*, 3, 60-74. <https://doi.org/10.18384/2224-0209-2020-3-1025>
- Martyshkin, A. I., & Pashchenko, D. V. (2019). *Modeli i sredstva sbora i predstavleniya informacii iz otkrytyh social'nyh setej dlya ispol'zovaniya v sistemah prinyatiya reshenij [Models and tools for collecting and presenting information from open social networks for use in decision-making systems]*. Penz; Publishing House. state technol. University. <https://expose.gpntbsib.ru/expose/vnp-acf1bed4/book/%D0%932019-20419777670036>
- Novikov, A. Yu., Lukyanov, V. V., Dronik, V. N., Podolsky, E. V., & Novgorodtsev, A. M. (2018). Informacionnaya sistema podderzhki retrospektivnogo i tekushchego analiza emocional'noj tonal'nosti vydelennogo media-kontenta [Information system supporting retrospective and current analysis of the emotional tone of the selected media content]. *Scientific thought*, 2(28), 48-55. <https://vure.mil.ru/upload/site50/gO2MPvGN4i.pdf>
- Osipov, V., Osipova, M., Kuleshov, S., Zaytseva, A., & Aksenov A. (2022). Epidemiological informing of the population in cities: models and their application. *Foresight and STI Governance*, 16(2), 80-89. <https://foresight-journal.hse.ru/data/2022/06/22/1847484104/7-Osipov-80-89.pdf>
- Petrov, N., Arkhipova, A., Spiridonov, V., & Peigin, B. (2020). Infodemiya: sushchestvuyushchie podhody k analizu panik, fobij, sluhov, fejkov vo vremya epidemij i predlozheniya po bor'be s nimi [Infodemic: existing approaches to the analysis of panics, phobias, rumors, fake during epidemics and suggestions to combat them]. *Monitoring the economic situation in Russia: trends*

and challenges of socio-economic development, 8(110), 70-77.

<https://www.iep.ru/upload/iblock/96c/9.pdf>

- Rubtsova, Y. (2012). Avtomaticheskoe postroenie i analiz korpusa korotkih tekstov (postov mikroblov) dlya zadachi razrabotki i trenirovki tonovogo klassifikatora [Automatic construction and analysis of a corpus of short texts (microblogging posts) for the task of developing and training a tone classifier]. *Knowledge Engineering and Semantic Web Technologies*, 1, 109-116. https://study.mokoron.com/wp-content/uploads/2014/02/rubtsova-kesw_paper-21.pdf
- Smetanin, S. (2020). The Applications of Sentiment Analysis for Russian Language Texts: Current Challenges and Future Perspectives. *IEEE Access*, 8, 110693-110719. <https://doi.org/10.1109/ACCESS.2020.3002215>
- Vasilkova, V. V., & Legostaeva, N. I. (2020). Social'nye boty v komp'yuternoj propagande: serfing na informacionnoj volne koronavirusa [Social bots in computer propaganda: surfing the information wave of the coronavirus]. *Monitoring public opinion: economic and social changes*, 6(160), 329-356. <https://doi.org/10.14515/monitoring.2020.6.1762>