

*A. A. IMANBERDI<sup>1</sup>, L. L. LA<sup>2</sup>, R. N. MOLDASHEVA<sup>2\*</sup>*

*Eurasian National University named after L.N. Gumilyov, Astana, Kazakhstan*

*E-mail: asel\_khas@list.ru, lira\_la@mail.ru;*

*<sup>2</sup> Kazakh Agrotechnical University named after S. Seifullin, Astana, Kazakhstan*

*E-mail: raushan85\_07@mail.ru*

## **APPROACHES TO THE DESIGN OF A SOFTWARE SYSTEM FOR MONITORING AND ANALYZING DATA OF SOCIAL NETWORKS, ADAPTED TO THE SPECIFICITY OF THE KAZAKH AND RUSSIAN LANGUAGES**

*This article is a description of the process of designing a software system for data analysis and monitoring of social networks, adapted to the specifics of the Kazakh and Russian languages. Also, a review of existing systems used by ordinary users, government agencies, commercial organizations, as well as scientific organizations, and designed for the analysis and monitoring of social networks, was carried out.*

**Key words:** *social network, social network monitoring, social well-being, user perception assessment, software system.*

**Introduction.** Today, the Internet has become the main means of communication, and social networks have become the main platform for communication and interaction. Moreover, social networks are a source of huge amounts of information and are widely used for research purposes, marketing, various business tasks, tasks related to security at various levels, recruiting agencies, etc. The use of such sources is not an easy process, and the problems that arise along the way require specialized technologies and tools. The main tasks of software systems for analyzing data from social networks are: achieving an understanding of the processes occurring in social networks (analysis and monitoring), as well as transferring the social network to the required state (forecasting and management).

Existing software tools used in the activities of companies in our country have a number of shortcomings. Basically, such developments are foreign, the linguistic dictionaries of which are poorly adapted to the specifics of the Russian and Kazakh languages; any situation is monitored manually, there is no practice of using Data Mining methods to process the information array of streams.

As a result, the task of creating a software system for data analysis and monitoring of social networks becomes relevant. The implementation of the software tool will allow monitoring social networks and web resources, solving a wide range of business tasks in real time: identifying reputational and informational threats, criticism, negativity, misinformation, tracking the mood of the company's customers and responding to their complaints and suggestions, fighting fraud, brand management, formation of new sales channels, etc.

**Materials and methods of the study.** Review of Existing Systems for Monitoring and Analyzing Social Media Data. To date, there are quite a few tools for monitoring social

---

\* E-mail корреспондирующего автора: raushan85\_07@mail.ru

networks and content analysis. The foreign market is represented by such systems as TweetDeck, Hootsuite, PolyAnalyst, Socialmention and other analogues that allow using social networks to solve a wide range of business problems. As for the market of Kazakhstan, it is not richly represented by domestic developments in opinion monitoring. There are such systems as iMAS, Alem Media Monitoring. Table 1 presents the functionality of software systems.

**Table 1** – Functionality of software products

Compare parameter	Sales percentage, %				
	Proposed system	Alem Media Monitoring	iMAS	PolyAnalyst	IQBuzz
Russian language support	100	100	100	0	100
Kazakh language support	100	100	100	0	0
English language support	100	0	100	100	0
Modeling social well-being	100	0	0	0	100
Analysis detail	100	0	0	0	0
Social media monitoring	100	100	100	100	100

As shown in the table, the systems have a number of disadvantages. In the considered software systems, there are no modules for managing the tone dictionary for modeling social well-being. All of the above systems have a number of disadvantages. Foreign systems do not take into account the syntax of Russian and Kazakh languages and are aimed mainly at foreign users. They, among other things, have a rather high cost, and free versions, in turn, do not provide access to all the functionality of the program and algorithms for finding information, which significantly reduces the efficiency for finding answers and user comments.

The proposed software system for monitoring and analyzing data from social networks will be implemented based on the use of the functionality of similar systems, and is primarily aimed at research purposes, taking into account the syntax of the Russian and Kazakh text, and will also be more accessible to users. The system will use: a semantic profile model, machine learning algorithms, production model methods, which will make it possible to build rules for assessing social behavior, and to perform a quantitative analysis of the tone of the text.

**The body. Outcomes.** *Software system design concept. Functional Software Tool Design.* The context diagram of the software system model is shown in Figure 1.

The input is information converted by the function block. For this model, the input information is social network data. The governing mechanism is the regulations of social networks, dictionaries. The executing mechanisms are the system itself and the moderators of the system. Various types of reports on public opinion, quantitative reports on the assessment of the tonality of topics with monitoring visualization serve as output information .

The software system consists of three stages: «Data entry», «Data processing», «Report generation».

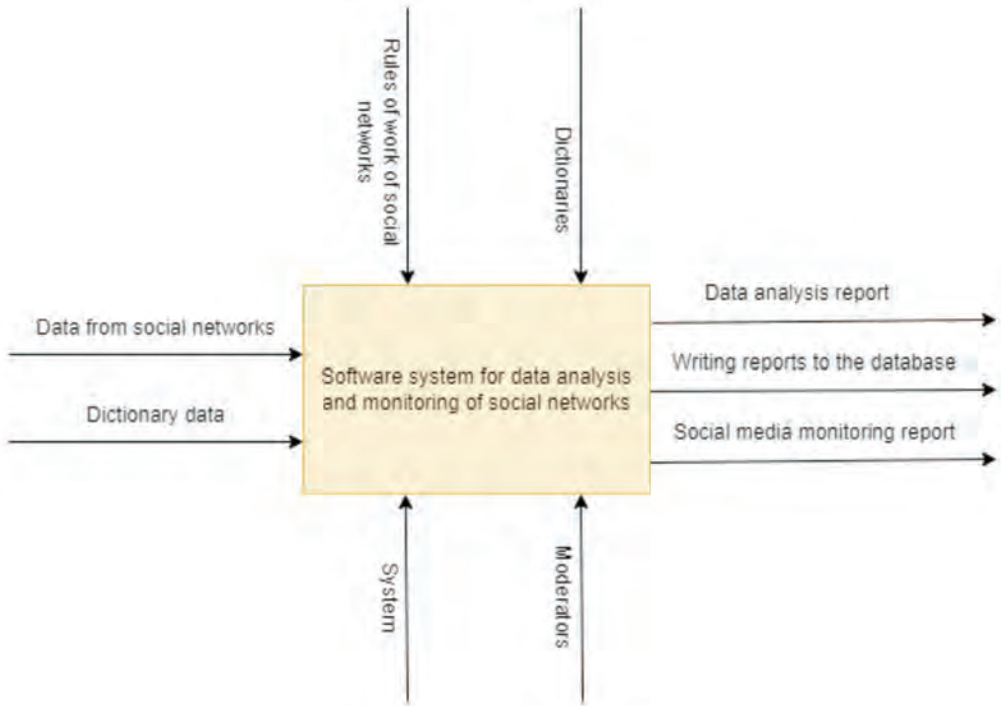


Figure 1 – Context diagram of the software system model

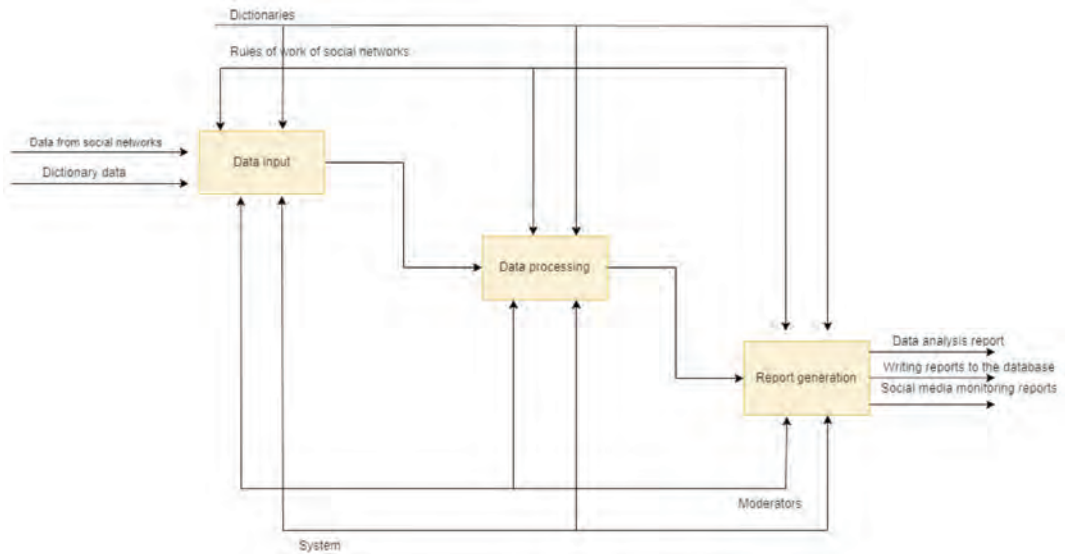
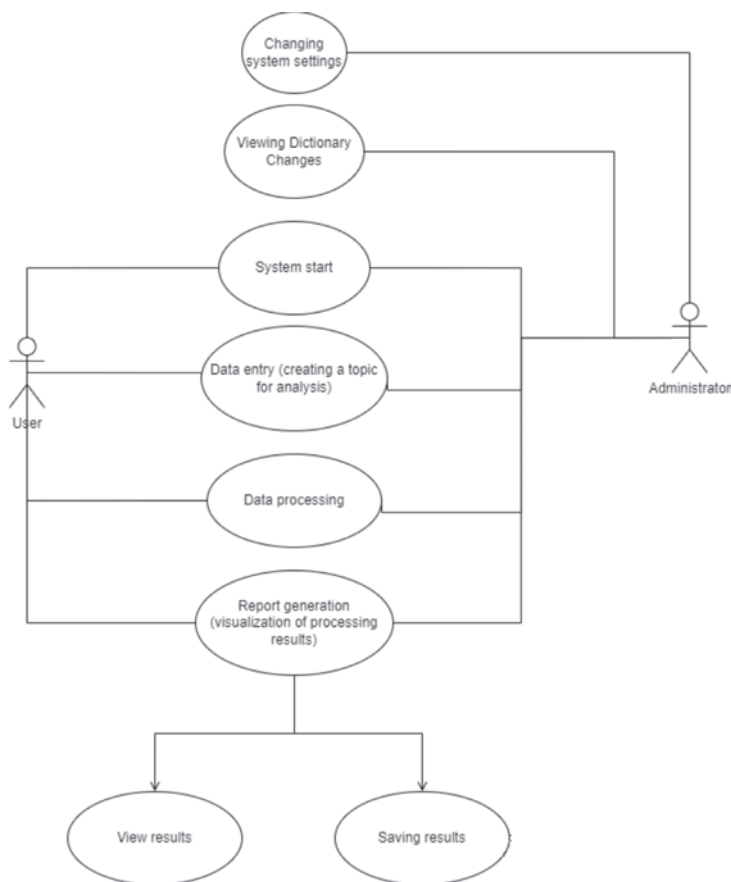


Figure 2 – Decomposition diagram of the first level

Next, we will simulate the problem using UML. Figure 3 is a system use case diagram



**Figure 3** – Usage diagram

The system must use a database that will contain all the necessary data tables. Figure 4 shows an entity-relationship diagram, which is a data model and includes entities and relationships that reflect the main business rules of the subject area.

The architecture of the designed system is based on a three-level model of a client-server application (DBMS - graphical user interface - business logic).

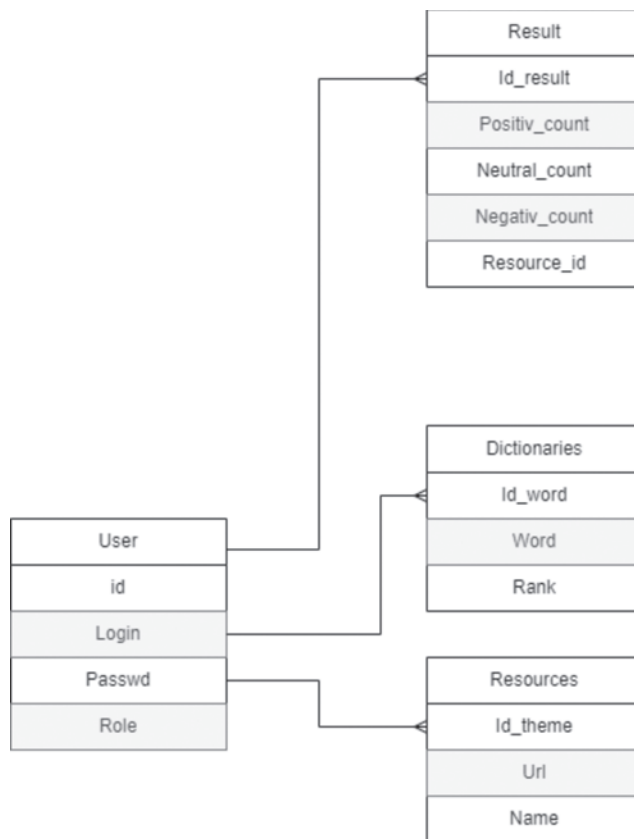
**Software system requirements.** *Functional requirements.* *The system must:*

- Have the “Connector” module that allows you to configure connection settings to Internet resources and APIs of well-known social networks from the resource catalog;

- Have a module «Linguistic constructor» that allows you to create a tonal dictionary in Kazakh and Russian;

- Have the «Data Processing and Analysis Platform» modules that allow modeling social well-being and building a quantitative analysis of the results of monitoring social networks, processing data in order to determine the assessment of sentiment, using machine learning algorithms;

- Have a module «Visualization generating reports» in PDF format and allowing you to build graphical representations of the results of quantitative analysis;



*Figure 4 – Entity-relationship diagram*

– The System should be able to respond to public publications or comments that are displayed as a result of monitoring in a special section, without entering the specified social networks. Also, from the public comments (mentions) displayed in the monitoring results, it should be possible to create an internal task;

– The System should be able to collect feedback left on the company's mobile application, hosted by the company in the Google Play /App Store application store. In the System, in addition to collecting such comments, it should be possible to respond to them without entering the Google Play / App Store application store.

- Work with texts in 2 languages (Kazakh, Russian);
- Determine the tone of the topic/text (negative, positive, neutral);
- Identify the profile of a participant in a social network by reading data on the participant's profile; counting the participant's activity in the topic (the number of comments, likes, reposts);
- Predict the behavior of society members, model social interaction;
- Calculate the index of social well-being of society.

*Integration requirements:*

The system must have built-in modules for connecting to the API of social networks Vkontakte, Facebook, Twitter, as well as to the API of mobile applications in the Google

Play and App Store. Monitoring in social networks should be carried out according to the following rules:

1. Vkontakte:

- words or phrases;
- links;
- hashtags (prefixed with “#”)
- mention of any user/community page (with the “@” prefix);
- comments in the discussion on the personal page of the community;
- comments on the personal page of the community.

2. Twitter:

- words or phrases;
- links;
- hashtags (prefixed with “#”);
- mention of any page of the user (with the prefix “@”);
- tweets / publications / posts on the personal page of the user/community account.

3. Facebook:

- mentioning the user/community page in public publications/posts;
- mentioning the user/community page in public comments;
- comments on the personal page of the user/community.

*Non-functional requirements (characteristics of the program and its environment). The system must have the following characteristics:*

- be scalable. This is an obvious requirement for almost all large data processing systems.

As the volume of data increases, the speed of the system should increase;

- must be expandable. The number of data sources on the Internet is constantly increasing. In addition, as the number of analytical tasks grows, so does the number of types of data that need to be collected. Therefore, the system must be extended with handlers for new data types;

- there must be a storage system. This is necessary primarily to ensure that the same data is not collected multiple times.

*User interface requirements:*

- the appearance of the user interface and forms of interaction with the user should be intuitive;

- the user interface should provide access to the internal functionality of the system, to various digital content;

- the presence of a mobile version.

*Information system security requirements:*

- The system must have reliable means of protecting data from unauthorized access;

- system should have a security system at the level of best practices. The functionality of the security system should provide: user identification; verification of the user's authority when working with the system; differentiation of user access at the level of tasks and information arrays in accordance with job responsibilities; blocking password guessing at all entry points to the system

- the safety of information in the system must be ensured in the event of the following emergency situations: – power failure, – software failure, – hardware failure, – database destruction;

– The system must have reliable means of protecting data from unauthorized access;  
– the system must be operational 24 hours 7 days a week, downtime - no more than 20%.

**Conclusion.** Despite the fact that the foreign market is represented by a sufficient number of tools for monitoring social networks and content analysis, there is a need for such a system in Kazakhstan. There is a need to identify consumer opinions, identify reputational and informational threats, criticism, negativity, disinformation, monitor the mood of the company's customers and respond to their complaints and suggestions, combat fraud, brand management, the formation of new distribution channels, etc. The projected system will allow developing a new direction the market for software solutions for monitoring and analyzing the opinions of social networks in order to obtain information on assessing user perception of content and assessing the social well-being of society.

## REFERENCES

- 1 Tabassum S, Pereira FSF, Fernandes S, Gama J. Social network analysis: An overview. WIREs Data Mining Knowl Discov. 2018;e1256. <https://doi.org/10.1002/widm.1256A>.
- 2 Zhurba A.K., Prodan E.A., Sushkova M.S. Information systems for data analysis: formalized data analysis from social networks // international scientific journal “innovative science” No. 5/2016 ISSN 2410-6070. pp. 69-72.
- 3 Zhang, B. and Vos, M. (2014), “Social media monitoring: aims, methods, and challenges for international companies”, Corporate Communications: An International Journal, Vol. 19 no. 4, pp. 371-383. <https://doi.org/10.1108/CCIJ-07-2013-0044>.
- 4 Top 10: Social media monitoring services. - URL: <https://www.cloudcms.ru/tools/smm/> (accessed 02/24/2022).
- 5 Social Media Monitoring Tools. - URL: <https://www.trustradius.com/social-media-monitoring/> (Accessed 02/25/2022).
- 6 Electronic fund of legal and normative-technical documentation: official site. – URL: <http://docs.cntd.ru/> (date of access: 10/11/2020). – Text : electronic.
- 7 Rubtsova Yu. V. Methods and algorithms for building information systems for classifying texts of social networks by tone: specialty 05.13.17 “Theoretical foundations of computer science”: dissertation for the degree of candidate of pedagogical sciences / Rubtsova Yulia Vladimirovna; Institute of Informatics Systems. – Novosibirsk, 2019.
- 8 Bazhenov N. I., Gubanov D. A. Review of information systems for the analysis of social networks / Management of large systems. Issue 41. M.: IPU RAN, 2013. P. 357-394.
- 9 Kuznetsova, M. M. Features of automating the monitoring of social networks / M. M. Kuznetsova. - Text: direct // Young scientist. - 2020. - No. 26 (316). - S. 33-36.
- 10 Kulikov S.Yu. Automatic extraction of opinions: linguistic aspect: specialty 10.02.21 - Applied and mathematical linguistics: dissertation for the degree of candidate of philological sciences / Sergey Yurievich Kulikov; Institute of Linguistics RAS. - Moscow, 2016. - 200 p. – Bibliography: p. 39–64.

***А. А. ИМАНБЕРДИ<sup>1</sup>, Л. Л. ЛА<sup>1</sup>, Р. Н. МОЛДАШЕВА<sup>2</sup>***

*<sup>1</sup>Л. Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан  
<sup>2</sup>С. Сейфуллин атындағы Қазақ агротехникалық университеті, Астана, Қазақстан*

## ҚАЗАҚ ЖӘНЕ ОРЫС ТІЛДЕРІНІҢ ЕРЕКШЕЛІГІНЕ БЕЙІМДЕНГЕН ӘЛЕУМЕТТІК ЖЕЛІЛЕРДІҢ ДЕРЕКТЕРІН БАҚЫЛАУ ЖӘНЕ ТАЛДАУ БАҒДАРЛАМАЛЫҚ ЖҮЙЕСІН ЖОБАЛАУ ТӘСІЛДЕРІ

Бұл мақалада қазақ және орыс тілдерінің ерекшеліктеріне бейімделген әлеуметтік желілердегі деректерді талдау және мониторингілеуге арналған бағдарламалық жүйені жобалау үдерісінің сипаттамасы берілген. Сондай-ақ, қарапайым пайдаланушылар, мемлекеттік органдар, коммерциялық ұйымдар, сондай-ақ ғылыми ұйымдар пайдаланатын және әлеуметтік желілерді талдау мен мониторингілеуге арналған қолданыстағы жүйелерге шолу жасалды.

**Түйін сөздер:** әлеуметтік желі, әлеуметтік желі мониторингі, әлеуметтік әл-ауқат, пайдаланушы қабылдауын бағалау, бағдарламалық қамтамасыз ету жүйесі.

**А. А. ИМАНБЕРДЫ<sup>1</sup>, Л. Л. ЛА<sup>1</sup>, Р. Н. МОЛДАШЕВА<sup>2</sup>**

<sup>1</sup>Евразийский национальный университет им.Л.Н.Гумилева, Астана, Казахстан

<sup>2</sup>Казахский агротехнический университет им. С. Сейфуллина, Астана, Казахстан

## ПОДХОДЫ К ПРОЕКТИРОВАНИЮ ПРОГРАММНОЙ СИСТЕМЫ МОНИТОРИНГА И АНАЛИЗА ДАННЫХ СОЦИАЛЬНЫХ СЕТЕЙ, АДАПТИРОВАННОГО ПОД СПЕЦИФИКУ КАЗАХСКОГО И РУССКОГО ЯЗЫКА

Данная статья представляет собой описание процесса проектирования программной системы анализа данных и мониторинга социальных сетей, адаптированного под специфику казахского и русского языка. Также проведен обзор существующих систем, применяемых обычными пользователями, государственными органами, коммерческими организациями, а также научными организациями, и предназначенных для анализа и мониторинга социальных сетей.

**Ключевые слова:** социальная сеть, мониторинг социальных сетей, социальное самочувствие, оценка пользовательского восприятия, программная система.