International Journal of Psychosocial Rehabilitation, Vol. 24, Issue 08, 2020
ISSN: 1475-7192

Approaches to the Examination of the Practice of Monitoring "Big Data" of a Socio-Economic Nature

[1]Aleksandra Polyakova

Abstract*--The article considers approaches to the examining of the practice of monitoring "big data" of a socio-economic nature. Big data, which is an array of structured and unstructured data that is difficult to process using traditional methods, has significant potential, allowing you to get online rapid diagnostic results, examine the entire data set, rather than samples, and use various machine algorithms that identify implicit relationships. The author of the article considers social scoring as promising by marketers, specialists in personnel search and bailiffs, collection and detective agencies for searching debtors in social networks, collecting data, and identifying relationships. It should be noted that many are not ready for data inspection, despite the fact that the automation of business processes in this area has been active for more than 15 years in Russia. For example, in the field of HR Analytics, this is due to the fact that a relatively small number of companies have implemented ERP systems that allow one to accumulate the necessary data, as well as the HR specialists themselves do not save all the data.*

Key words*--monitoring, "big data", social information, designing of monitoring systems.*

I. INTRODUCTION

Practical algorithms run within the monitoring system to further design of analytical forms and adaptation of most appropriate algorithms for the collection and processing of information will lead to recommendations on the regulation of social and economic policy, as well as for the development of projects of state programs in the medium and long-term in the relevant fields of regulation. Currently, most of the algorithms are developed in the natural Sciences, but the experience of using them for large data sets of socio-economic nature and distributed computing is extremely small. The results of the study will allow us to get closer to the justification of the possibility of application and to the choice of ways to adapt existing techniques. The basis of the assessments will be a combination of traditional methods of analysis, the tools and methods of which are reflected in a large number of works, with experimental methods.

There is potential for further adaptation of the system to meet specific challenges in the most sensitive areas of National Projects. Such as identifying and excluding the causes of child suicide, deviant behavior, extremist tendencies, etc. based on implicit links and monitoring cascades of data in social networks in the absence of direct statistical observation and obvious signs indicating the presence of corresponding inclinations.

II. MATERIALS AND METHODS

Problems of formation and implementation of socio-economic policy were considered by such researchers as X. Bayer, A. Wagner, G. Geller, V. Sombart, B. Turner, and S. Hill. The essence of modern socio-

[1]*Industrial University of Tyumen, Volodarskogo Street, 38, Tyumen, Russian Federation, Email ID: agpolyakova@mail.ru*

DOI: 10.37200/IJPR/V24I8/PR280132
Received: 21 Jan 2020 | Revised: 08 Feb 2020 | Accepted: 14 Mar 2020

economic policy is revealed in the writings of J. G. Baldock, S. Wickerstaff, E. Giddens, K. Deutschman, G. Esping-Andersen, and a number of other researchers. A special role in understanding the mechanisms of social policy implementation was played by the works of L. Erhard (the concept of the social state). Institutional aspects of socio-economic development are presented in the works Of F. Blok, H. Demceza, P. DiMaggio, O. Williamson, Th. Schumpeter etc. However, among the numerous works, we would like to mention a recent work by N. Taleb, focused on a fundamentally new paradigm for ensuring the stability of systems based on dynamic disequilibrium. Approbation of the ideas presented by N. Taleb will be possible when applying to monitoring systems with large data sets.

A significant contribution to the work on the problem of design of monitoring systems, control and support of managerial decisions such domestic and foreign scientists as Ansoff I. H., L. A. Bazilevich, V. N. Burkov, P. Drucker, A. Y. Kabanov, B. Z. Milner, G. Mintzberg and others made. Issues of performance management systems such authors as J. V. Bogatin, M. B. Miller, F. M. Rusinov, A. M. Smolkin, I. M. Syroezhkin, O. G. Turovets, and etc. examined.

III. RESULTS

Today, there is a process of building up the government's own information and analytical infrastructure [1]. This is due to the development of our own platforms, as well as access to successful sociological services that implement the experience of foreign universities. As a result, the focus is made on the developing our own research sociological on-line panels, as well as a system for collecting expert information.

At the same time, resources are aimed not only at collecting data, but also at processing them attracting increasing interest, which allows us to obtain high-quality results and identify implicit interdependencies.

Computer modeling of social mechanisms has had more than 60 years of history. It is used for various purposes – for building scenarios for the development of processes, checking the consistency of descriptive theories, examining the mechanisms of development of emerging phenomena, etc. As a rule, mechanistic computational models form a common basis for the social sciences.

The use of digital technologies in analysis has always been associated with the simulation of processes or the search for social mechanisms. Even before computer modeling, there were mathematical models for this purpose that were focused on identifying social influence, the influence of the structure of social groups on the information flow, and identifying the fundamental properties of social networks[2]. Recently, most of the work has been focused on computer modeling. Now it can be considered as the final stage in monitoring socio-economic processes, allowing you to reconfigure the existing information cascade. For example, in the works of J. Coleman, and John and Jin Gullahorns the prospects of computer simulation in sociology are considered. J. Coleman considered questions about social actions and social organization, as well as the possibility of using simulation to test social smoking scenarios among adolescents [3]. John and Jin Gullahorns studied the behavioral science of organization and conflict resolution. [4] R. McGinnis presented a stochastic model of social mobility, in which he considered "mobility" as a change in the position of an individual in any sociometric observable system (including physical space) [5].

As a rule, earlier computer modeling is used to test conceptual models aimed at understanding interpersonal communication and behavior of individuals. Computer programming forced researchers to break

DOI: 10.37200/IJPR/V24I8/PR280132
Received: 21 Jan 2020 | Revised: 08 Feb 2020 | Accepted: 14 Mar 2020

down social phenomena into algorithmic blocks that help identify social mechanisms. In addition, modeling in the social sciences contributed to the examination of poorly understood situations and phenomena that are not available for experiment, including the representation of human cognitive processes. For example, testing of Twitter data showed that online social networks exhibit many features typical for social systems, with highly grouped individuals in a topology without scale [6]. Twitter data was used to test the theory of the theoretical cognitive limit of the number of stable social relationships, and it was found that people tend to have common connections in the same metropolis, and the distance between them, borders, and language differences affect their relationships [7]. Many examinations have focused on determining which and how information is transmitted over the network [8] and understanding the mechanisms of information dissemination – for example, as in the case of viral content - to identify influential distributors and understand their role [9].

Today, computational methods used to test the completeness of the theoretical foundations (concept modeling) or to detect a hypothesis are becoming more common [10]. A number of works have focused on the examining of social dynamics during the emergence of protests, and evidence of social influence and viral spread can serve as a tool for empirically testing the mechanisms outlined theoretically in the framework of collective action models.

Many research papers are separate case studies based on large data sets generated over a period for a specific problem. M. Bier and K. Wagner [11] disclose features of monitoring business processes based on the use of social media. The areas of anti-crisis management are considered by S. Hiltz, P. Diaz and D. Mark [12], as well as by S. Stieglitz, D. Bunker, M. Mirbabae and K. Enis [13], policies by S. Stieglitz and L. Dang-Suan [13], in the field of entertainment – by Th. Shen, K. Hok Chuan, and S. Cheng [14].

One of the reasons for the popularity of social networks is the ability to receive or generate information events and public messages, as well as share them everywhere with low costs and without spatial restrictions.

Despite the variety of subjects, they have a lot in common. In particular, the algorithm for obtaining the required information or knowledge from social networks is similar. In many ways, this determines the development of such a direction as "social network analytics", which is focused on combining, expanding and adapting methods for examining social network data. Despite its development, there is still insufficient information on social media inspection capabilities, as well as on general models and approaches. S. Aral, S. Dellarocas, and D. Godes developed the basics of organizing research in social networks [15], and V. van Osh and S. Kursaris proposed a structure and research program [16] for working with organizational social networks. Both options focus on the classification of research areas, as well as not solving research questions, rather than on methods for solving them.

Recently, social media has become a source of receiving and distributing information more and more often. They have found their application in a variety of fields and most often, the experience of their application is revealed in the foreign literature. The growing involvement of the population in social networks opens up new opportunities for examining different communication models. For example, social media data can be used to identify problems in the social environment, trends, influential actors, and other types of information. S. Golder and M. Mesi monitored the behavior of individuals using Twitter data and considered how people's moods change over time of the day, day of the week, and season [17].

DOI: 10.37200/IJPR/V24I8/PR280132
Received: 21 Jan 2020 | Revised: 08 Feb 2020 | Accepted: 14 Mar 2020

A separate layer of work is devoted to solving social problems and examining related events based on big data arrays. They reflect research methods, techniques, and algorithms that allow researchers to implement them to solve problems that affect both individuals and society as a whole. Solutions to such problems can be provided by measuring public opinion and identifying signs of harmful behavior through predictive inspection. In particular, in the works of U. Kursuncu, M. Gaur, U. Locale, K. Thirunaryan, A. Sheth, and I. B. Arpinar methods are shown for detecting and predicting such following phenomena [18]:

Harassment, which manifests itself as an aggressive exchange of words that leads to emotional experiences and stress, leaving social networks, and then out of life. According to a Pew Research Center survey, more than 70% of Internet users over the age of 18 have experienced harassment, and 40% have been subjected to it. Of the latter, two-thirds were related to social networks. In addition, according to a report by the cyberbullying research center, 25% of teenagers said they were being humiliated online.

Criminal acts and violence with the use of weapons. Criminal networks take advantage of social media to put pressure on competitors, and the identification of such users helps law enforcement to anticipate crime before it can occur. For example, the work of L. Balasuriya, S. Viyeratne, D. Doran and A. Shet reflects the problem of searching for members of street gangs in Twitter [19]. In their study, nearly 400 profiles of gang members were identified manually using initial terms, including rappers associated with the gang, their retwitters and followers using text tweet functions, descriptions and comments on YouTube videos, emoticons, and image profiles to work with various machine learning algorithms, including Bayesian algorithm, logistic regression, and so on. Such solutions can be used to assist law enforcement officials through awareness-raising, as well as predicting conflicts between groups for possible incidents involving weapons. Similar algorithms are also used to detect cyberbullying, violence among young people, etc.

Special attention should be paid to the work with the "collective mind", which is considered by researchers as a separate area of work. Users are characterized by special feelings, ways of expressing them, decisions, intelligence and share their opinions in networks that form a public view. The diversity of data can be seen as a force of "collective intelligence" that can be used to form recommendations, judgments, and strategies. In addition, it is a well-known fact that an intelligent verdict made by a crowd is superior to an individual decision. The examination of public opinion is necessary for making decisions and minimizing private biases that can reduce the objectivity of decisions, combine different points of view and knowledge, and improve the coverage and completeness of the analysis. However, none of the existing works illustrates the concept of "collective mind" or public opinion as a holistic phenomenon statistically and analytically. Today, most of the population uses social networks, and a methodologically verified way of measuring the diversity of the crowd is crucial for increasing the social activity of people in social networks and the evolution of the social system [20].

One of the areas of monitoring that is of significant interest is related to the spread of the "infection effect", which is based on the formation of information cascades. In this case, the term "infection" refers more to the transmission of data and broadcasting of information messages. The "contagion effect" has been studied in the framework of viral marketing: for example, 79.2% of forum participants are known to help a friend make a decision to buy a product, 47.6% of them are non-project participants, and 65% of forum participants share tips both offline and in person, based on information they have read on the Internet.

DOI: 10.37200/IJPR/V24I8/PR280132
Received: 21 Jan 2020 | Revised: 08 Feb 2020 | Accepted: 14 Mar 2020

Although "infestation" is not considered in the literal sense, it is interesting to note that social networks, especially those such as Twitter and Facebook, play a significant role in disaster management, spreading information about emergencies among members of the disaster-affected community.

The latter is considered the fourth most popular source of access to emergency information. Many works have examined social media data to understand networks and extract critical information to develop a pre-and post-disaster mitigation plan. In this way, analysis allows you to transform social network data sources into knowledge. For example, parishes in Louisiana actively used their social networks to share information with the community affected by the disaster (for example, a flood map, emergency shelter locations, medical services, and garbage disposal) during 2016 Louisiana flood that damaged more than 60,000 homes. Researchers are generally interested in the patterns generated by the cumulative interaction of online users on Facebook during a disaster response, which allows for a deeper understanding of the role of using social networks to spread information about emergencies. The results of the examination showed that monitoring systems of this type consist of three groups of actors: individuals, emergency agencies, and organizations. The core of the social network includes many individuals of people. They actively participate in the exchange and updating of information. Emergency agencies and organizations are located on the periphery of the social network, connecting the community with other communities.

When examining promising solutions and analytical procedures, it should be noted that they should be based on a wider range of information sources, including social networks, electronic media resources, search engine data, as well as specialized information resources, as well as integrated with speech and image recognition modules. The accumulated information arrays can be replenished both in a constant mode and at the user's request. In this case, data collection should be carried out on a streaming basis with the possibility of dynamically joining or disconnecting information search areas. The amount of information processed determines the use of a distributed storage with a decentralized data catalog, which will allow one to search for previously collected information and process it using intellectual means, taking into account the morphology, chronology, and additional attributes of the object.

One of the promising results of the message text examining can be the completion of the glossary and clarification of the parameters for filtering out the noise background through morphological examination and synthesis of word forms using alternative modules from different developers. Additional functionality includes the following items:

- filtering out duplicate and irrelevant information items;
- identification of an array of proper names, taking into account the possible use of transliterations or replacement of Cyrillic letters with similar Latin letters;
- automatic categorization of information flows by the extended classifier;
- formation of thematic clusters for building private algorithms for their research;
- identification and clarification of objects of interest through linguistic tools that may not be identified in advance;
- building a dynamic object profile that allows you to interactively visualize and provide the necessary information both through the search interface and through intelligent linking to a previously formed context.

DOI: 10.37200/IJPR/V24I8/PR280132
Received: 21 Jan 2020 | Revised: 08 Feb 2020 | Accepted: 14 Mar 2020

Integration with geo-information systems and services that take into account the geolocation of objects is a significant direction in the development of the functionality of the designed system. This can enrich the information and analytical resources of the management decision-making system by possible examination of movements based on the secondary data, spatial dynamics of public opinion, accurate linking, systematization and integration with other features of the system. This will allow one to model processes in dynamics and in the future, taking into account the spatial factor [21].

This functionality will be able to integrate geo-spatial characteristics of the object and its satellites for the purpose of drawing on an online map and correlating with the results of text data analysis. The user functionality in this case should include the following:

- localization of the object and points of interest to it, other objects associated with the examination;
- segmentation of analytical procedures by geography (regions, groups of regions, manually defined areas);
- filtering out irrelevant objects based on the specified time limits or other characteristics;
- superimposing an ego network (social subnet) on a geographical map for clustering and identifying hotspots of tension and risk accumulation;
- use of analytical visualization tools (color schemes, quantitative characteristics, etc.);
- preparation and translation of analytical reports into presentation format, including interactive reports.

Additional functionality can provide the ability to notify promptly interested parties and significant facts and events, such as changes in the information background or the occurrence of grounds for disciplinary or other penalties.

In any case, automated typing, in which a multivariate typological model is built that shows the characteristic properties of an individual, should be considered as one of the promising technologies that is becoming more widespread [22]. For example, in 2018 Sberbank announced the introduction of a system for scoring borrowers in accordance with psychometric models, including using the so-called digital footprint, which is data collected when using social networks, mobile traffic, and, potentially, photos posted on the Internet. The company Mail.ru Group identified features that would characterize certain social groups without using survey methods for each individual, for example, uses automated typing. Thus, automated typing technology is used not so much to evaluate an individual user, but to identify the portrait characteristics of user groups, which can play a role in planning and managing changes.

Company specialists Mail.ru the authors noted that the hypothesis put forward the possibility of obtaining data describing the psychological characteristics and personal preferences of an individual without individual poll was confirmed. For such examination, the company used solutions from the open source segment, which allowed modeling and processing large amounts of data. Among other things, the company's internal developments (Tarantool) were used, as well as those that are generally recognized among IT specialists in Data Science, such as Hadoop and Spark. It is also worth noting the latest Machine Learning algorithms that are implemented in the Scikit –learn libraries, Spark ML, XGboost, Vowpal Wabbit, Tensorflow.

Automated typing is also characterized by the ability to reduce costs in the implementation of personnel procedures-both in terms of cost and time characteristics. However, the most obvious economic effect can be found when working with typical positions, and not when searching for employees to implement non-standard

DOI: 10.37200/IJPR/V24I8/PR280132
Received: 21 Jan 2020 | Revised: 08 Feb 2020 | Accepted: 14 Mar 2020

projects. Searching for specific actors involves using a different model that is specifically designed to target individuals in the network and identify actors that match the specified characteristics [23]. Based on this model, special audience segments were built, after which advertising campaigns were created within myTarget and Vkontakte, which are targeted at the identified audience segments. When using this technology, for example, when recruiting staff, the time for closing vacancies in certain positions was reduced by half, and the cost was reduced by four times. A similar solution can be used to achieve other goals related to the search for actors with specified characteristics, or as a component of a risk management system to restrict individual access to information cascades.

In this context, it should be noted that the monitoring system should have a multi-level security system that allows the protection of personal data, confidentiality of the collected information, reliable storage of information, as well as an appropriate level of protection from erroneous and malicious actions of system users due to such features as the differentiation of access rights depending on the level of the user's management hierarchy, exclusion of access through unprotected data transfer protocols, logging of actions and events in the context of the user, storing protocols and providing administrator access to view them through a separate interface.

IV. DISCUSSION

Research papers summarizing practical experience in building monitoring systems published by international organizations are of significant interest. For example, since there is no single indicator for assessing an individual's socio-economic status, many attempts are being made to develop a composite index to measure it. The experience of the Statistical office of the European Union, which provides data for comparing European countries and regions, on the use of summary indicators can be considered representative [24; 25]. They are an integral part of the infrastructure created and maintained by Eurostat and are defined as a summary measure related to a key issue or phenomenon derived from a series of observable facts. Summary indicators can be used to display relative positions or to display positive / negative dynamics. In contrast to the data from which they are derived, indicators are used for specific purposes, which are determined by the questions it is intended to answer and the parties who ask these questions [26].

The project on examining public opinion of the School of public administration was known very well. J. Kennedy of Harvard University, specializes in studying the political views of American youth, as well as the work of the National center for public opinion research at the University of Chicago, which conducts research for such structures as the Pentagon, the us Department of energy, the us Bureau of statistics, NASA, and World Bank.

Practice shows that the experience of using big data arrays is not so extensive, although it is becoming more common. Big data has been widely used in the management of various objects, from corporations to public authorities. Among the leaders in the use of big data at the government level is China, whose State Council recently announced guidelines for simulating the development of big data technologies. Thus, at the end of 2017, the development of a platform for the exchange of interdepartmental data was completed. In addition, by the end of 2018, a single database for exchanging government data should be operational. It is expected that big data technologies will contribute to efficient use of resources, improve management, introduce industrial innovations, create new business models, and support economic restructuring in general.

DOI: 10.37200/IJPR/V24I8/PR280132
Received: 21 Jan 2020 | Revised: 08 Feb 2020 | Accepted: 14 Mar 2020

Monitoring systems developed by individual companies are usually among the scoring systems. Almost every resident is connected to social networks that provide a lot of data. Companies use platforms to sift through data collected from Facebook, Google+, LinkedIn, or Twitter. These platforms collect data (such as keywords used by candidates in posts) and turn it into usable information. Such practice is extremely diverse. Among the pioneers in the field of working with social network data include Alfa-Bank, HomeCredit, Tinkoff credit systems, OTP Bank, Sociohub.ru, Double Data, Skorista analytical Agency, IQ'men, Palantir, Social network agency. Analytical reviews of these organizations are available, as well as publications of the Analytical center under the government of the Russian Federation, which summarizes the experience of various organizations in line with the examined issues.

A whole block of developments related to the problem under examination is related to the development of social scoring systems. Social scoring ideas have been developed and widely used by Klout. Using data from social networks, a "Klout Score" is generated, which is taken into account by many American companies. For example, the Palms Hotel in Las Vegas provides benefits for its guests based on their Klout score. And Virgin Airlines offers free tickets on its new flights to customers with a high social score.

Social scoring is a type of scoring that evaluates individuals based on their social characteristics and predicts their behavior by inspecting their presence in social networks. Based on metadata such as gender, age, place of residence, position, duration of work in one place, education, etc., an individual is assigned certain weights. PwC's developments on credit scoring point the use of social media data such as friends and contacts, unversity, specialty, moves and "checks-in", stop words, marital status, profile age, frequency of publications, updates, children and relatives, consumer and social portrait, foreign language proficiency, life events, subscriptions, communities, groups, photos, hobbies, change history, search for additional profiles, and semantic profile examination. In total, according to PwC, you can upload up to 300 characteristics from social networks, in contrast to the standard fixed data with no more than 40 parameters. Next, each new individual is evaluated from the position of its compliance with the specified characteristics, i.e. it is automatically assigned a rank indicating the degree of trust and attention to it. As a rule, credit organizations use scoring models that extract data from social networks.

If we analyze the use of social scoring, the most notable is the experience of China, where the idea came to fruition in 2007, when the first developments of the social credit system were initiated. The prototype of that was the system for assessing the creditworthiness of borrowers of the American company FICO-a pioneer in the use of intelligent analytics and data science on improving operational solutions that ensure the security of transactions of billions of payment cards, timely return of leased transport, etc. In China, the "program for creating a social credit system (2014-2020)" was announced, significantly changing the original concept. The new idea was that by 2020, all business entities – business units and citizens-should be included in the system of such assessment in real time and receive a rating, and data on this rating should be open and accessible to everyone (for example, on the CreditChina portal). As a result, holders of a high rating will be able to enjoy various benefits, while holders of a low rating will have trouble, ranging from price discrimination, ending with difficulties with employment, etc., thus implementing the principle of social justice.

The system is already being tested in a number of cities. Thus, residents of Rongcheng city received a starting rating of 1000 points, which increase or decrease depending on the behavior of a citizen based on the

DOI: 10.37200/IJPR/V24I8/PR280132
Received: 21 Jan 2020 | Revised: 08 Feb 2020 | Accepted: 14 Mar 2020

information received in real time from various sources: municipal, commercial, law enforcement, judicial and other authorities – in a single center, where it is processed using big data technology. The array of estimated information includes more than 150 thousand different parameters, data about which is received in a single cloud. It is administered and processed by the eight largest private companies. The data includes not only formally collected information, but also information about consumer behavior obtained from mobile applications, which give operators information about where, when, what goods and services were purchased and in what quantity, what people are interested in, what their social circle is, and so on. For example, Alibaba, the operator of the Sesame Credit service, evaluates such parameters as the authenticity of the name and other personal data specified during registration in social networks, the rating of a person's contacts in social networks, the time spent on online games, the volume of purchases of goods for children, and the number of p2p loans issued or received using online services. Chinese big data initiatives do not stop at social scoring.

The Chinese government has announced the creation of the international Big Data Science Center, designed to:

(1) fill in technological gaps and create a unique interdisciplinary cloud service platform for collecting and processing GEODATA about resources, ecology, and the environment;
(2) support important decisions in the field of science and technology and promote breakthrough scientific discoveries in the field of science and technology infrastructure;
(3) build a data-based, high-performance digital Earth science platform that promotes the concept of sustainable development.

It is expected that this project will provide high-quality support for the creation of the digital economy in China and the implementation of the national big data strategy. It is emphasized that this project should give way to the introduction of innovative projects, the development of institutional mechanisms designed to promote openness and data exchange, and ensure the timely transformation of scientific and technological achievements into information and technological decision-making platforms.

Social scoring in the field of credit based on the use of big data arrays has a long history: the American company Big Data Scoring has been using social scoring in the credit sphere for more than 30 years. Companies that lend to borrowers based on information obtained only from social networks without meetings with potential borrowers (LendUp, Wonga) operate successfully. For example, Kreditech has a cloud service that provides automatic customer scoring based on an expanded array of big data, which received $ 4 million investments from Bloomberg Capital. Using the project Lenddo.com (Philippines, Colombia) you can get a loan using your social media profile.

In last 4-5 years, scoring has been gaining popularity in Russia. For example, the National credit Bureau also plans to check the accounts of borrowers in social networks. The new services are based on Double Data technologies and are based on publicly available sources of information-social networks. The Social Link solution is a service that enriches information about the borrower with data from publicly available information sources collected in a single user-friendly interface. Social Link data combines contact information, information about the borrower's education and employment (current and previous), photos, lists of close related persons, and more. The main use of the Social Link service is manual verification of the borrower's personal data and underwriting of the loan application. NBCH provides a technical platform, a service channel, and provides all operational support

DOI: 10.37200/IJPR/V24I8/PR280132
Received: 21 Jan 2020 | Revised: 08 Feb 2020 | Accepted: 14 Mar 2020

related to the operation of new solutions. Social Attributes and Social Link data are received within the existing NBCH infrastructure, which saves NBCH customers from integration costs and allows them to use the innovation within a few days after making a decision to connect to the service.

Some of the pioneer organizations that use social network-based big data analysis in their work, taking into account the development stage and / or the subject of their interest, are listed below.

- Alfa-Bank, HomeCredit-testing the scoring model;
- Tinkoff credit systems-analyzes how long the user has been registered, the number of friends, and activity;
- OTP Bank checks whether the client has friends with a positive credit history;
- Sociohub.ru performs scoring for banks and microfinance organizations;
- Double Data offers solutions to reduce the level of overdue debt;
- Skorista analytical Agency performs scoring, semantic content analysis;
- IQ'men automates processes and performs analytical processing of information;
- Palantir performs analytical processing of big data;
- Social network agency is engaged in complex reputation management and marketing.

Within the Russian market, there is a Crediograph solution, which is a social scoring model based on a web service. The platform's basis is a self-learning algorithm for assessing the borrower's solvency, which already uses Facebook-based assessment. LinkedIn, VKontakte, Foursquare.

Social scoring is seen as promising by marketers, search specialists and bailiffs, collection and detective agencies to find debtors in social networks, collect data, and identify relationships. It should be noted that many are not ready for data examination, despite the fact that the automation of business processes in this area has been active for more than 15 years in Russia. For example, in the field of HR Analytics, this is due to the fact that a relatively small number of companies have implemented ERP systems that allow you to accumulate the necessary data, as well as the HR specialists themselves do not save all the data.

V. CONCLUSION

In the course of the examination, data sets were formed that characterize the socio-economic situation of the population in sections that are either not collected by the Federal state statistics service and other organizations, or are aggregated very non-operatively. This will allow us to achieve the following in the future:

- to fill in technological gaps and create a unique interdisciplinary platform of services for collecting and processing data on income and expenses of the population, the state of the labor market, changes in the migration situation, the social well-being of the population and changes in its financial behavior strategies;
- to provide support for making important decisions in the field of socio-economic situation of the population and promote breakthrough scientific discoveries in the field of scientific and technical infrastructure;
- to build a high-performance digital resource based on big data that contributes to operational monitoring of the socio-economic situation of the region's population.

The implemented research can serve as a basis for the practical implementation of digital sociology. It should allow the introduction of innovative mechanisms designed to promote openness and data exchange, and

DOI: 10.37200/IJPR/V24I8/PR280132
Received: 21 Jan 2020 | Revised: 08 Feb 2020 | Accepted: 14 Mar 2020

ensure the timely transformation of scientific and technological achievements into information technology platforms to support the process of development and management decision-making.

Promising areas of work include the following:

- designing the architecture of an innovative information and analytical system for collecting and processing large socio-economic data extracted from the Internet space;
- searching and adapting of algorithms for collecting, processing, filtering and examining large data sets that characterize the socio-economic situation and identify the reactions of the population in the region;
- developing and descripting of reporting forms for analytical work and design of the system architecture, including the potential for today's functionality for the use of big data technology of a socio-economic nature.
- forming the concept of using big data and monitoring the socio-economic situation of the population and promptly tracking its reactions to current events and proposed changes in the region;
- developing of recommendations for the formation and applying of a system for monitoring the socio-economic situation and reactions of the population in the Internet space;
- designing and descripting the procedures for conducting analytics within the developed monitoring system;
- substantiating of the list of proposals for integrating the proposed system into the infrastructure of information support for the processes of developing and making managerial decisions in the field of social and economic policy.

Financing

The study was funded by RFBR and Kurgan oblast according to the research project № 19-411-450001/19

REFERENCES

1. Kibakin, M.V. (2019) Using Internet technologies for social diagnostics of public confidence in the police. *Social and humanitarian technologies*. 2, 18-24.
2. Edling, C.R. (2002) Mathematics in sociology. *Annu Rev Sociol*, 28, 197-220. DOI: 10.1146/annurev.soc.28.110601.140942
3. Coleman, J.S. (1965) The use of electronic computers in the study of social organization. *Eur J Sociol*, 6, 89-107. DOI: 10.1017/S0003975600001156
4. Gullahorn, J.T, Gullahorn, J.E. (1963) A computer model of elementary social behavior. *Syst Res*, 8, 354-362. DOI: 10.1002/bs.3830080410
5. McGinnis, R. A stochastic model of social mobility. Am SociolRev. (1968) 33:712–22. doi: 10.2307/2092882
6. Kwak H, Lee C, Park H, Moon S. What is twitter, a social network or a news media? In: Proceedings of the 19th International Conference on World Wide Web. Raleigh, NC: ACM (2010). pp. 591–600.
7. Takhteyev Y, Gruzd A, Wellman B. Geography of twitter networks. Soc Netw. (2012) 34:73–81. doi: 10.1016/j.socnet.2011.05.006
8. Java A, Song X, Finin T, Tseng B. Why we twitter: understanding microblogging usage and communities. In: Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis. San Jose, CA: ACM (2007). pp. 56–65/
9. Bakshy E, Hofman JM, Mason WA, Watts DJ. Everyone's an influencer: quantifying influence on twitter. In: Proceedings of the Fourth ACM International Conference on Web Search and Data Mining. Kowloon: ACM (2011). pp. 65–74
10. Holme P., Liljeros F. Mechanistic models in computational social science. Front. Phys., 17 September (2015). doi.org/10.3389/fphy.2015.00078
11. Beier, M., Wagner, K. (2016). Social media adoption: barriers to the strategic use ofsocial media in SMEs.Proceedings of the european conference on information systems.

12. Hiltz, S.R., Diaz, P., Mark, G. (2011). Social media and collaborative systems for crisismanagement, ACM Transactions on Computer-Human Interaction, Vol. 18 Issue 4, December 2011, ACM New York, NY, USA, DOI:https://doi.org/10.1145/2063231.S.
13. Stieglitz, S., Bunker, D., Mirbabaie, M., Ehnis, C. (2017a). Sense-Making in Social MediaDuring Extreme Events. Journal of Contingencies and Crisis Management (JCCM).http://dx.doi.org/10.1111/1468-5973.12193.
14. Shen, Y., Hock Chuan, C., Cheng, S. H. (2016). The Medium Matters: Effects on WhatConsumers Talk about Regarding Movie Trailers. In Proceedings of the InternationalConference on Information Systems. Simmonds,
15. Aral, S., Dellarocas, C., Godes, D. (2013). Introduction to the special Issue—Socialmedia and business transformation: a framework for research.Information SystemsResearch, 24(1), 3–13. http://dx.doi.org/10.1287/isre.1120.0470.
16. van Osch, W., Coursaris, C. K. (2013). Organizational Social Media: A Comprehensive Framework and Research Agenda. In 2013 46th Hawaii International Conference on System Sciences (HICSS). p. 700–707. DOI: 10.1109/HICSS.2013.439
17. Golder, S. A., Macy, M. W. (2011). Diurnal and Seasonal Mood Vary with Work, Sleep, and Daylength Across Diverse Cultures. Science, 333(6051), 1878—1881.https://doi.org/10.1126/science.1202775
18. Kursuncu U., Gaur M., Lokala U., Thirunarayan K., Sheth A., Arpinar I. B.. (2018). Predictive Analysis on Twitter: Techniques and Applications. arXiv:1806.02377v1 [cs.SI] 6 Jun 2018
19. Balasuriya L., Wijeratne S., Doran D., and Sheth A., Finding Street Gang Members on Twitter, in ASONAM, 2016.
20. Gerasimov, V.O., Sharafutdinov, R.I., Kolmakov, V.V. et al. (2019). Control in the human capital management system in the strategy of innovative development of a region. Entrepreneurship and Sustainability Issues, 7(2), 1074-1088. DOI: 10.9770/jesi.2019.7.2(20)
21. Polyakova, A., Kolmakov, V., Yamova, O. (2019). Regional competitiveness response to innovation changes: Issues of evaluation. Journal of Urban and Regional Analysis, 11(2), 159-172. DOI: 10.37043/JURA.2019.11.2.3
22. Zavyalova, N. et al. (2019). Dataset on the interview-based survey of Moscow bicycle infrastructure. Data in Brief, 26 DOI: 10.1016/j.dib.2019.104429
23. Tkachenko, E., Rogova, E., Bodrunov, S., Dmitriev, N. (2019). Valuation of intellectual capital in the context of economic potential of A company. Paper presented at the Proceedings of the European Conference on Intellectual Capital, 2019-May, 303-314.
24. Zavyalov, D. V., Saginova, O. V., & Zavyalova, N. B. (2017). The concept of managing the agro-industrial cluster development. Journal of Environmental Management and Tourism, 8(7), 1427-1441. doi:10.14505/jemt.v8.7(23).12
25. Polyakova, A. G., Ramakrishna, S. A., Kolmakov, V. V., & Zavyalov, D. V. (2019). A model of fuel and energy sector contribution to economic growth. International Journal of Energy Economics and Policy, 9(5), 25-31. doi:10.32479/ijeep.7849
26. Gagarina, G.Y., Dzyuba, E.I., Gubarev, R.V., Fayzullin, F.S. (2017). Forecasting of Socio-Economic Development of the Russian Regions. Ekonomika regiona – Economy of region, 13(4), 1080-1094.
27. Banchiamlak Nigussie Tefera, Young-Dong Kim, and . 2019. Ethnobotanical study of Medicinal plants used as Antimalarial and Repellent by Sidama people of Hawassa Zuria district, Southern Ethiopia.. *Journal of Complementary Medicine Research*, 10 (1), 13-26. doi:10.5455/jcmr.20181102063241
28. Sharma, R., Balhara, Y.P.S., Sagar, R., Deepak, K., Mehta, M.Heart rate variability study of childhood anxiety disorders(2011) Journal of Cardiovascular Disease Research, 2 (2), pp. 115-122. DOI: 10.4103/0975-3583.83040

DOI: 10.37200/IJPR/V24I8/PR280132
Received: 21 Jan 2020 | Revised: 08 Feb 2020 | Accepted: 14 Mar 2020