International Journal of Economics and Business Administration
Volume VII, Special Issue 1, 2019

pp. 26-39

System Data Analysis: Innovative Technologies, Methods and Techniques

Submitted 02/01/19, 1st revision 15/02/19, 2nd rervision 21/04/19, accepted 15/05/19

Garnov, A.[1], Zvyagin, L.[2], Sviridova, O.[3]

Abstract:

Purpose: *The purpose of this article is to explore the use of technology of system data analysis. To achieve this goal it is necessary to solve the following set of tasks: to consider the essential characteristics of system analysis and its concept and to identify the algorithm of the application system of data analysis in the framework of methodological tools.*
Design/Methodology/Approach: *The system analysis of data is considered a popular method of scientific knowledge, similar to the study of operations, since this approach is aimed at optimizing the components of a particular system. Its proper application needs knowledge of its methods, technologies and techniques. Authors compiled, analyzed and structured all the elements needed to solve the system data analysis' issues.*
Findings: *Authors developed methodological tools, which can be classified according to some features as: the stage of solving a specific problem, the method of obtaining and presenting information data, the level of formalization of data. Authors also presented the system approach algorithm and scope, highlighting the technologies involved.*
Practical Implications: *The results stressed the problem that the state is faced with the need to maintain a balance between the desire to regulate the state in this sphere and the needs of the market. The findings may be implemented into managerial practices in order to improve and increase the effectiveness of business administration.*
Originality/Value: *The key contribution of this study is the compiled set of data analysis elements to practically keep business environment indicators of interest, relying not only on intuition, but also on framework of methodological tools.*

Keywords: *Data analysis, system analysis, big data, information modeling, decision making.*

JEL codes: *C55, C61.*

Paper Type: *Research article in Special Issue dedicated to Russian Economy.*

Section 1: *Digital Economy.*

[1]*Plekhanov Russian University of Economics, Moscow, profgarnov@yandex.ru*
[2]*Financial University under the Government of the Russian Federation, Moscow, saineco@yandex.ru*
[3]*Financial University under the Government of the Russian Federation, Moscow, olshan@list.ru*

1. Introduction

According to the scientific definition, data analysis is a field of science that deals with the construction and research of mathematical methods and computational methods for extracting knowledge from existing data, as well as the process of selection, transformation and modeling of data in order to obtain the necessary information and management decisions. Analysis is closely related to modeling. A model is an object or description of an object that, under certain conditions, can replace the original, reproducing the properties and characteristics of the original. Models can be empirical, theoretical, and mixed. The basis of empirical models are the methods of sensory cognition as: observation; experiment; measurement and comparison.

Theoretical models are obtained with the help of mathematical laws, based on rational methods of knowledge such as: analysis; synthesis; classification; abstraction; formalization; analogy; deduction; induction, etc.

Theoretical models are built with the help of empirical facts. For example, induction is a logical theoretical operation of movement from a particular fact to a general position. However, verification of each particular fact is required, which means that it is based on empirical knowledge, not theoretical. And to build an empirical model requires a preliminary theoretical basis - the choice of the object for observation, hypothesis detection, etc.

Therefore, theoretical and empirical models complement each other, forming a mixed model – that is a model based on the previous two types of models. Building models, modeling facilitate the study of objects and systems of interest, allow you to detect dependencies, make predictions. The main thing is that the model well and accurately reproduces the functioning of the simulated system. The most informative model of the system is the "black box" model. It is a system in the form of a rectangle with an internal structure unknown to the researcher and a set of input and output variables (effects).

The process of building a model begins with the definition of the problem, the problem situation. For example, you want to know which customers make the most profit for your company? Or what products are bought and sold together? In other words, the task is to create a model of associations and forecast profitability.

The next step in the construction of the model is the systematization of data, the selection of the most important factors that will significantly affect the final result. The larger the data model, the more complex it is considered.

The third step is to find a model that explains the source data. The model can be known, formal, or exceptional, based on the experience and intuition of the analyst. Task is reduced to search to a method for the solution of an objective. If you want to

build a sales forecast model, you can use the average method by entering an additional seasonality factor for the accuracy of the measurements.

In the fourth stage, the resulting model is tested for quality in practice. If the quality of the model is acceptable, it is put into operation. Otherwise, either factors are added or excluded, or the model itself is changed. The main principles of building models are:

- ✓ consider the expertise of the expert;
- ✓ study the problem from different perspectives, to combine methods;
- ✓ move from simple models to complex and more accurate models;
- ✓ check the model periodicaly for accuracy, use new information, data etc.;

Thus, the system analysis of data is a popular method of scientific knowledge, similar to the study of operations, since this approach is aimed at optimizing the components of a particular system.

2. Literature Review

Data are phenomena, facts or processes in the form of raw measurement results and observations. Data that is used for analysis are:

- ✓ numerical (number of workers, 100 rubles);
- ✓ interval (20%, market share);
- ✓ nominal (sex, age, profession of the consumer);
- ✓ rank (more/less, better/worse).

Data can be compared with "new", while their analysis - with the ability to "clean" and realize their potential. Data analysis is a discipline dealing with decision-making problems in an environment where the choice of an alternative requires the analysis of complex information of different physical nature (Rupeika-Apoga *et al.*, 2018). The subject of data analysis are the following:

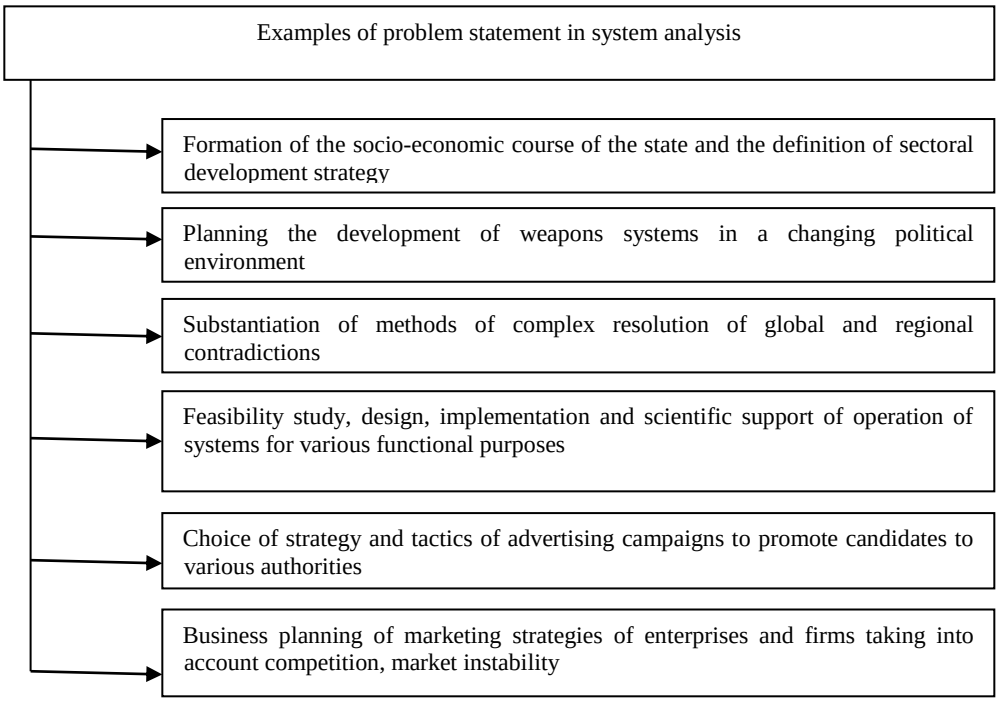
- ✓ theoretical studies aimed at finding optimal solutions to the problems through a systematic approach;
- ✓ some research on selected aspects of interrelations of the components of the system;
- ✓ integration of methodological tools, which is developed in the process of system analysis or other areas.

In addition to the model, there are three important components in data analysis: expert, analyst and hypothesis. An expert is a specialist who professionally solves the problems of a particular subject area. The expert puts forward a hypothesis, tests the sample data to build models. Thus, in combination of computer programs and expert participation, truly effective analytical solutions are obtained.

Analyst-a specialist in the field of tools and software for data analysis, for example, Data Mining (Polyakova *et al.,* 2019). In addition, it generates expert opinions, tests hypotheses and analyzes the results.

Hypothesis-the assumption put forward by the analyst and expert on the issue of interest. The object of data analysis is the problems that reflect the formation of new and existing systems from a conceptual, technological, information, economic point of view. Examples of such problems are shown in Figure 1.

Figure 1. *Problems to solve with system analysis (Compiled by the authors).*



Since the implementation of the system analysis is aimed at solving the problems presented in Figure 1, such an approach solves a number of the following problems:

- ✓ identification and streamlining of structured goals of the relevant system, potential ways to achieve them;
- ✓ modeling of system functionality;
- ✓ search and selection of criteria for comparative evaluation of decisions;
- ✓ division of a particular system into elements and identification of the close relationship between them;
- ✓ determination of the importance of a separate chain of criteria, models, activities, technologies.

It should also be said that data analysis is a set of the following key procedures:

- ✓ formation of goals in the development of the system;
- ✓ study of the components of the system, a designation of its structure through analysis of the relationships;
- ✓ algorithmization in achieving the goal;
- ✓ model construction and verification.

Data analysis is a set of techniques and methods from the fields of statistics, mathematics, natural sciences and economics in order to verify the truth of assumptions and hypotheses, scientifically based interpretation of the studied facts.

3. Results

Data analysis has several similar concepts due to the fact that it covers a variety of approaches and methods in various fields of science and activity. First of all, we are talking about such concepts as artificial intelligence (AI), machine learning and previously mentioned Data Mining. The term artificial intelligence is a scientific direction for solving problems of software or hardware modeling of intelligent human activities. In other words, an automatic AI system is able to perform creative functions of a person. Artificial intelligence is divided into applied and strong AI. The task of applied AI is to solve with the help of computers and software such creative tasks that do not go beyond the cognitive human ability. For example, speech recognition, computer vision, handwriting, etc. At the present stage, a promising direction is a strong artificial intelligence, which should not be inferior to the human intellect and even surpass it. Machine learning is the acquisition of new knowledge based on algorithms and computer programs. Machine learning is the science that explores computer algorithms that improve during problem solving.

The systematic approach is aimed at the development of certain recommendations that contribute to the maintenance of the course of action with the active use of scientific achievements. Intensification of data analysis is more due to the fact that it has the ability to use computer technology that helps to use various methodological tools contained in the model. The search for the optimal solution of the problem is based on the formulation of the immediate goal of the system. However, there is a need to compare objectives and potential solutions to the problems, taking into account available resources and their needs. Here, data analysis has the ordering of actions in the study of the problem area with the use of methodological tools in the scientific areas.

System data analysis involves the use of methodological tools, which can be classified according to some features: the stage of solving a specific problem, the method of obtaining and presenting information data, the level of formalization of data (Figures 2-4).

Figure 2. *Methodological tools for the stage of solving the problem (Compiled by the authors)*

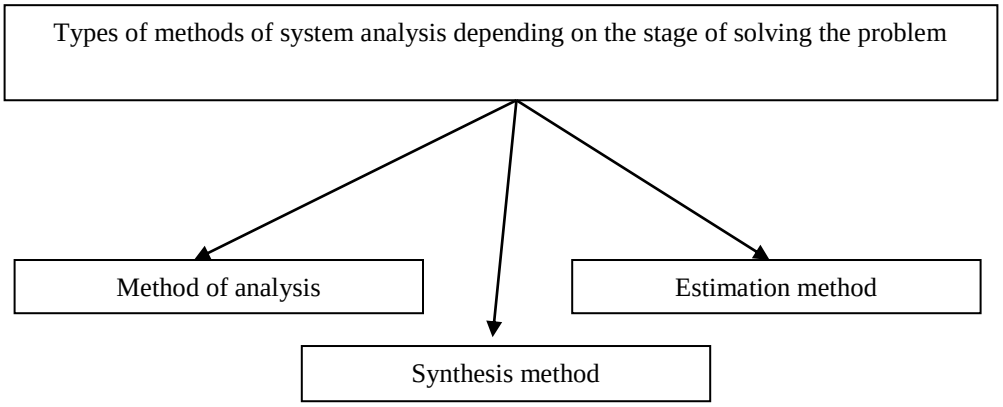


Figure 3. *Methodological tools for the method of obtaining information (Compiled by the authors)*

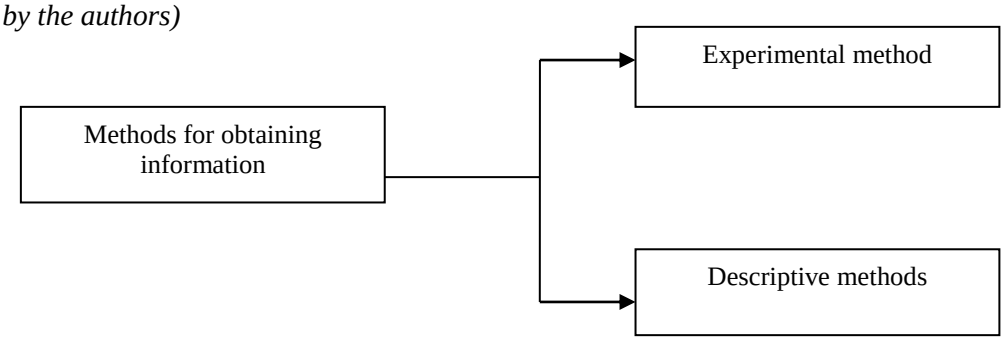
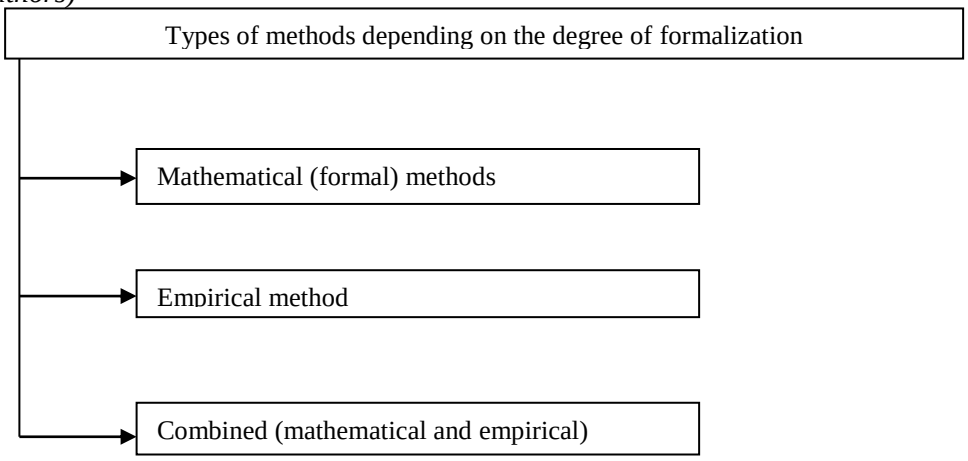


Figure 4. *Methodological tools on the level of formalization (Compiled by the authors)*



The system approach allows for a full and multi-faceted verification of alternatives through the evaluation of quantitative and qualitative indicators of costs and efficiency of the system. Model always investigates any method – quantitative and/or qualitative. Quantitative mathematical calculations provide an opportunity to assess the related decisions and to consider informal factors increasing the objectivity of the existing opinions of researchers. In recent years, the information method of analysis has became widespread, which makes it easier for analysts to understand and apply modern mathematical methods in modeling.

Methods of analysis are diverse. It all depends on what questions need to be answered. However, all methods are divided into two classes: one-dimensional and multidimensional. Multivariate methods allow you to simultaneously investigate the relationships of two or more variables and to test hypotheses about the causal relationships between them. Let's take a closer look at the key methods and examples of how they can be used.

Association:
This is the most well-known and available method of data analysis, which is a comparison of two or more elements to find a connection between them. It is used, in particular, in the field of retail trade in the layout of goods, sales promotion. Marketers have noticed that most often buyers together with strawberries buy cream, and together with pasta they buy sauce.

Classification:
This method is found in a variety of areas of human activity. It allows you to select homogeneous groups of data, objects by type, set of certain characteristics, as well as to have the data in a certain order, reflecting the degree of their similarity. It can be used to identify a particular class by comparing attributes with a known definition. Attributes are age, social group, gender. Examples of classification are the classification of goods to a certain product group, the classification of the client to one of the risk groups (to issue a loan/refuse to issue a loan).

Clustering (cluster analysis):
Clustering uses the common attributes in different classifications to identify clusters. The task of cluster analysis is to divide data into groups (clusters or classes) with similar values to make the right management decisions. For example, it is necessary to study the demand market, for this purpose consumers are divided into groups according to several groups of characteristics-preferences of price and quality of goods. It is also used in the classification of suppliers, customer segmentation in trade and finance, identifying situations in which there is a marriage in the workplace, what personal preferences of consumers bring the greatest profit, etc.

The difference between classification and clustering is the presence of predefined classes. It is only necessary to determine which class the object belongs to. While

cluster analysis is aimed at determining the presence or absence of relationships between objects.

Factor analysis:
This analysis allows us to study and measure the impact of factors on the value of the studied economic indicators. This accumulates the general patterns of data, and their features are discarded. There are several types of factor analysis such as:

- ✓ deterministic - the investigated indicator in the form of sum, product or particular factors;
- ✓ stochastic (correlation) - probabilistic relationship between the factors and the effective indicator;
- ✓ deductive - from general to particular;
- ✓ inductive – from particular to general;
- ✓ static and dynamic;
- ✓ retrospective and prospective.

Neural network:
These are models of biological neural networks of the brain, a "black box", as its input data is fed the output is the result. This method of automatic analysis is used in forecasting and adaptive control. With the help of neural networks, you can predict sales volumes, stock market indicators and other important information. The method of neural networks solves the problem of automating the extraction of necessary knowledge from a huge amount of data, which are usually unstructured.

Decision tree:
It is a method of automatically analyzing data in a sequential structure, where each object has a single node that provides a solution (in the form of "if ... then..."). The decision tree starts with a question that has two answers. Each answer leads to the next question, which also has two possible answers, and so on. Decision trees are used to describe and identify data, classification, numerical prediction, regression analysis.

Prediction:
This method is aimed at analyzing trends, predicting events based on experience, analyzing past events or data instances. For example, to predict future profits.

Combinations:
These methods are most often used together, complementing and clarifying each other. For example, decision trees are used to build classifications. And classification specified when using the method of clustering.

Processing with memorization:

As a rule, historical data are used for comparative analysis and forecasting. In this regard, many processes are automated in order to include them in the new model if necessary and to make a decision faster. Within the framework of the system approach, certain logical elements are also viewed, which are formed in a specific sequence due to the problem area of study. Figure 5 shows the system analysis algorithm. The modeling process itself within the framework of a system analysis occupies one of the 3 key areas of the considered approach (Figure 6).

Figure 5. *System approach algorithm* ***(****Compiled by the authors)*

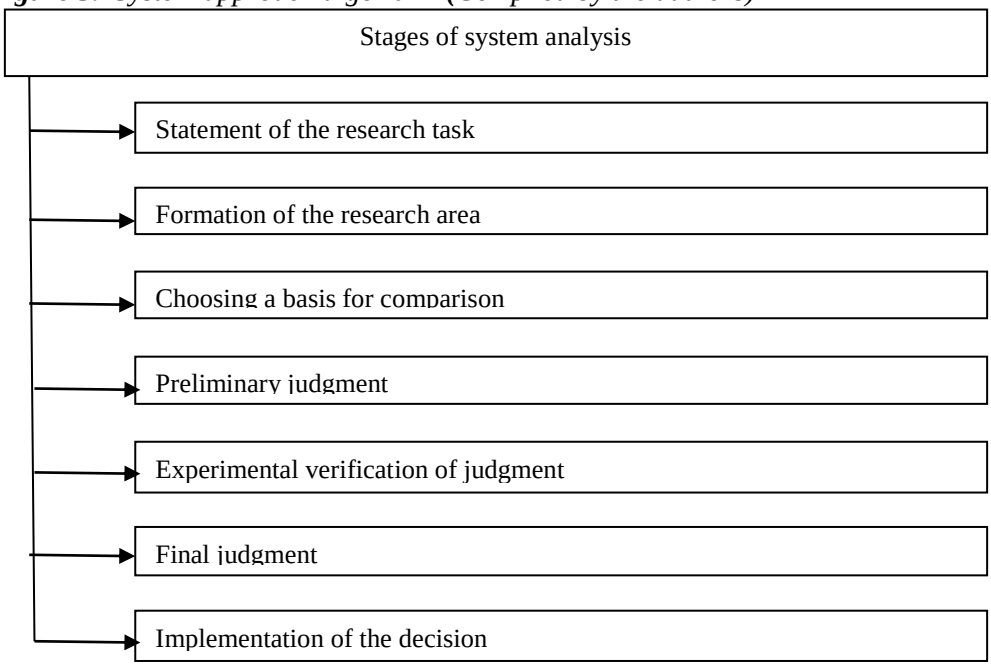
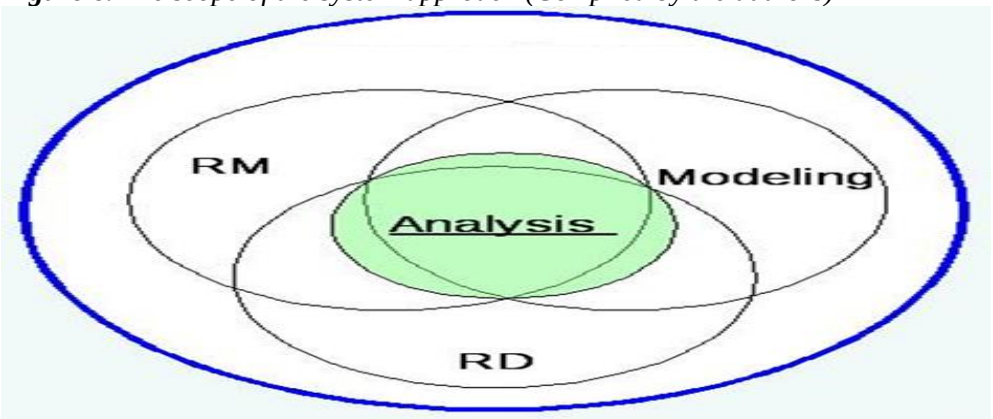


Figure 6. *The scope of the system approach (Compiled by the authors)*



In addition to modeling, it is also necessary to highlight the development and management requirements. All three areas in system analysis are interrelated and intertwined. Work on each area consists of a comprehensive analysis. Depending on the type of data used for analysis, the following technologies can be distinguished:

- ✓ Data Mining (mining);
- ✓ Text Mining (text analysis);
- ✓ Visual Mining (visual analysis);
- ✓ OLAP (online analysis);
- ✓ Process Analysis (Process Mining);
- ✓ Analysis of Web-resources (Web Mining);
- ✓ Real-time Analysis (Real-time data Mining).

Let's take a closer look at some of the data analysis technologies. The term Data Mining (translated into Russian means the detection of "machine" knowledge in primary data, data mining) studies the process of finding new and useful knowledge in databases by sifting a huge amount of "raw" material. By "machine" we mean algorithms, artificial intelligence. This technology is used for:

- classification of objects, events, observations;
- clustering-dividing the set of input vectors into groups of similar features;
- predictions;
- deviation analysis;
- compression of volumes necessary for the analysis of information.

In practice, Data Mining technology is used, to form targeted offers to customers based on their preferences. Thanks to the Internet Data Mining used by ordinary users every time they type their query in Google or other search engines. Text Mining analyzes text information by mathematical methods, interpolation, approximation, extrapolation, fuzzy methods. The tasks solved by Text Mining are reduced to obtaining structured information, classification and clustering of data, determining relationships, knowledge topics, and others. For example, this technology allows you to automatically classify into several groups published over a certain period of time by an Internet site lenta.ru articles and choose a topic of interest, for example, politics.

Web Mining-analysis of information extracted from Internet resources using special approaches and techniques that allow you to work with a special HTML markup language, which stores web documents. The practical application of the technology of Web Mining allows you to identify more sites similar to search web resources, identify visitor preferences when you use those or other resources on the Internet. Data analysis is a key step in designing a solution to the problem. With the help of data analysis, a model is created, which is further tested and accompanied by a potential solution.

Thus, the system data analysis is a set of methods and techniques underlying a particular algorithm of actions with the formation and use of the model in order to solve the problem. Modeling within the framework of system analysis is one of the most important areas along with the development and management of requirements. Work on each area consists of a timely and comprehensive analysis. Automated data analysis significantly reduces the time and cost of processing and implementing analytical projects.

4. Discussion

Today, organizations create huge amounts of data that are updated more and more often, and the outdated format of data storage and traditional methods of data analysis do not allow to use a huge layer of information to make meaningful research and conclusions. All this opens the way for big data. Big data is a deep and interactive work with a significant amount of information, which is often updated and is in different sources, in order to increase the efficiency of business, improve competitiveness. Internal big data sources: ERP, CRM, classifiers. External big data sources: social networks, Internet, specialized DataSet.

The term "big data" first appeared in 2008. Scientists continue to work on new methods of data analysis. In this work we have already got acquainted with some methods and technologies of big data analysis, which are borrowed from statistics and informatics (machine learning, classification, clustering, etc.). Every year CNews holds a conference "big data", where experts analyze the global market for big data analysis and predict its development. So, at the conference in 2018 was marked:

- ✓ systematic growth of the global market for big data business intelligence $122 billion in 2015, in 2016 it was130 billion;
- ✓ increased interest in big data analysis in Russia;
- ✓ the increasing demand of corporations in an IT company in bringing order to their data, systematization of reference data;
- ✓ focus on addressing global rather than local challenges when analyzing big data that affects the performance of the entire organization rather than its individual sites.

In April 2019, within the framework of the regular conference "Analytics and big data 2019: How to get an effective solution", specialists of leading foreign and Russian companies will consider the development of the big data market in 2018, its efficiency, impact on the development of other technologies, risks, solutions to personnel problems in the field of big data analysis, development prospects in 2019. It is already known that the first three places in the Top 10 data analysis technologies are "smart applications", "smart things" and “artificial intelligence”. More actively master the sphere of big data in the Telecom, public and financial sector, large industrial enterprises and energy.

A serious state task in this area was its regulation and security, suppression of attempts to use data analysis for other purposes with unhindered access to open data. With regard to social networks, the state is faced with the need to maintain a balance between the desire to regulate the state in this sphere and the needs of the market.

5. Conclusions

Currently, the system data analysis is a popular method of scientific knowledge, similar to the study of operations, because this approach is aimed at optimizing the components of a particular system. The key essential characteristics of the system data analysis are the object, the subject, the specification of the problem existing in the system. The systematic approach is aimed at the development of certain recommendations that contribute to the maintenance of the course of action with the active use of scientific achievements. The intensification of the development of system data analysis is more due to the fact that it has the ability to use computer technology that helps to use various methodological tools contained in the model.

Data analysis is a set of methods and techniques underlying a particular algorithm of actions with the formation and use of the model in order to solve the problem. Modeling within the framework of system analysis is one of the most important areas along with the development and management of requirements. Work on each area consists of a timely and comprehensive analysis. System analysis is a step in the design of the solution to the problem. With the help of data analysis, a model is created, which is further tested and accompanied by a potential solution. Data analysis is the process of research, filtering the data array to obtain the necessary and useful information and management decisions. Managers use data analysis to improve efficiency and grow their business. Huge amounts of data from companies are updated more and more quickly, which creates the prerequisites for leading big data technologies.

References:

Afsono, A., Agnello, A., Furceri, D. 2008. Fiscal policy responsiveness, persistence, and discretion. ECB Working paper no 954.

Alt, J.E., Lassen, D.D. 2006. Fiscal transparency, political parties, and debt in OECD countries. European Economic Review, 50(6), 1403-1439.

Alesina, A., Campante, F.R., Tabellini, G. 2008. Why is fiscal policy often procyclical? Journal of the european economic association, 6(5), 1006-1036.

Allegret, J.P., Raymond, H. and Rharrabti, H. 2016. The Impact of the Eurozone Crisis on European Banks Stocks, Contagion or Interdependence. European Research Studies Journal, 19(1), 129-147.

Azzimonti, M., Battaglini, M., Coate, S. 2016. The costs and benefits of balanced budget rules: Lessons from a political economy model of fiscal policy. Journal of Public Economics, 136, 45-61.

Andrikopoulos, A., Loizides, I., Prodromidis, K. 2004. Fiscal policy and political business cycles in the EU. European Journal of Political Economy, 20(1), 125-152.

Boldeanu, T.F., Tache, I. 2016. The Financial System of the EU and the Capital Markets Union. European Research Studies Journal, 19(1), 60-70.
Bouton, L., Lizzeri, A., Persico, N. 2017. The Political Economy of Debt and Entitlements https://drive.google.com/.le/d/0Bz2volNj-oH7SWlxX01zbm8zTlE/view.
Butters, M.J., Coate, S. 2008. A dynamic theory of public spending, taxation, and debt. The American Economic Review, 98(1), 201-236.
Becker, J.G., Gersbach, H., Grimm, O.R. 2010. Debt-sensitive majority rules. CEPR Discussion Paper, 7860.
Chen, C., Yao, S., Hu, P., Lin, Y. 2016. Optimal government investment and public debt in an economic growth model. China Economic Review.
Duguleana, L., Duguleana, C. 2016. Structural Aspects of the European Union Economy. European Research Studies Journal, 19(1), 93-128.
Fatás, A., Mihov, I. 2003. The case for restricting fiscal policy discretion. The Quarterly Journal of Economics, 118(4), 1419-1447.
Franzese, Jr.R.J. 2002. Electoral and partisan cycles in economic policies and outcomes. Annual Review of Political Science, 5(1), 369-421.
Fatás, A., Mihov, I. 2006. The macroeconomic effects of fiscal rules in the US states. Journal of public economics, 90(1), 101-117.
Hallerberg, M., Strauch, R. 2002. On the Cyclicality of fiscal policy in Europe. Empirica, 29, 183-207.
Henisz, W.J. 2004. Political institutions and policy volatility. Economics & Politics, 16(1), 1-27.
Gruner, H.P. 2017. Mechanisms for the Control of Fiscal Deficits. Journal of Economic Behavior and Organization, http://dx.doi.org/10.1016/j.jebo.2017.09.019
Lane, P.R. 2003. The cyclical behaviour of fiscal policy: evidence from the OECD. Journal of Public economics, 87(12), 2661-2675.
Lizzeri, A. 1999. Budget deficits and redistributive politics. The Review of Economic
Studies, 66(4), 909-928.
Leachman, L.L., Rosas, G., Lange, P., Bester, A. 2007. The political economy of budget deficits. Economics and politics, 19(3), 369-420.
Maltritz, D., Wüste, S. 2015. Determinants of budget deficits in Europe: The role and relations of fiscal rules, fiscal councils, creative accounting and the Euro. Economic Modelling, 48, 222-236.
Mink, M., De Haan, J. 2005. Has the Stability and Growth Pact Impeded Political Budget Cycles in the European Union? CESifo WP 1532 (2).
Mulas-Granados, C. 2003. The political and economic determinants of budgetary consolidation in Europe. European Political Economy Review, 1(1), 015-039.
Nordhaus, W.D. 1975. The political business cycle. The review of economic studies, 42(2), 169-190.
Person, T., Tabellini, G. 2001. Political institutions and policy outcomes: what are the stylized facts? CEPR Discussion Paper no 2872.
Perotti, R., Kontopoulos, Y. 2002. Fragmented fiscal policy. Journal of Public Economics, 86(2), 191-222.
Pontoh, W. 2017. The Capital Structure: Is Debt just a Policy or Requirement? European Research Studies Journal, 20(2A), 128-139.
Polyakova, A.G., Loginov, M.P., Serebrennikova, A.I. & Thalassinos, E.I. 2019. Design of a Socio-Economic Processes Monitoring System Based on Network Analysis and Big Data. International Journal of Economics & Business Administration, 7(1), 130-139.
Rupeika-Apoga, R., Zaidi, H.S., Thalassinos, E.Y. & Thalassinos, I.E. 2018. Bank Stability:

The Case of Nordic and Non-Nordic Banks in Latvia. International Journal of Economics & Business Administration, 6(2), 39-55.

Song, Z., Storesletten, K., Zilibotti, F. 2012. Rotten Parents and Disciplined Children: A Politico-Economic Theory of Public Expenditure and Debt. Econometrica, 80(6), 2785-2803.

Sørensen, B.E., Wu, L., Yosha, O. 2001. Output fluctuations and fiscal policy: US state and local governments 1978–1994. European Economic Review, 45(7), 1271-1310.

Volkerink, B., De Haan, J. 2001. Fragmented government effects on fiscal policy: New evidence. Public choice, 109(3), 221-242.

Ugurlu, E., Thalassinos, E., Muratoglu, Y. 2014. Modeling Volatility in the Stock Markets using GARCH Models: European Emerging Economies and Turkey. International Journal of Economics and Business Administration, 2(3), 72-87.

Wang, S., Yang, Z., Liu, H. 2011. Impact of urban economic openness on real estate prices: Evidence from thirty-five cities in China. China Economic Review, 22(1), 42-54.

Woo, J. 2003. Economic, political, and institutional determinants of public deficits. Journal of Public Economics, 87(3), 387-426.

Woo, J. 1999. Social Polarization, Economic Development, and Fiscal Instability: Theory and Evidence. Harvard University, Unpublished working paper.

Wu, G.L., Feng, Q., Li, P. 2015. Does local governments' budget deficit push up housing prices in China? China Economic Review, 35, 183-196.