Technical sciences

UDC 621.391

Kuzminykh Valeriy
Candidate of Technical Sciences, Associate Professor
National Technical University of Ukraine
"Igor Sikorsky Kyiv Polytechnic Institute"

Koval Oleksandr
Candidate of Technical Sciences, Associate Professor
National Technical University of Ukraine
"Igor Sikorsky Kyiv Polytechnic Institute"

Xu Beibei
Graduate Student of the
National Technical University of Ukraine
"Igor Sikorsky Kyiv Polytechnic Institute"

Zhu Shiwei
Graduate Student of the
National Technical University of Ukraine
"Igor Sikorsky Kyiv Polytechnic Institute"

MICROSERVICE ARCHITECTURE OF THE SYSTEM FOR ASSESSING THE LEVEL OF INTERNATIONAL ACTIVITY

Summary. *The article considers the structure of the adaptive request processing system, as developed in the framework of creating an information-analytical system for assessing the level of international activity. When building a request processing system, the use of microservice architecture is proposed. An extensive system of network hosts on which microservices are hosted, focused on the possibility of more efficient extraction of information from the* *source. The subsystem provides for the use of microservices as stand-alone containers, each of which is associated with specific information sources. The system provides for the adaptation of parameters, both individual microservices and script parameters. Docker and Kubernetes platforms are used to form and manage containers, which allows you to control the load on each host.*

Key words: *quality assessment, scenarios of analytical activity, information collection, microservice architecture, adaptation, big data.*

Introduction. One of the current trends in the progress of universities and research institutions at the national and international level is the forecasting of performance indicators, by which they are compared with similar institutions. This comparison is made by way of preparing the various ratings and in the form of monitoring the efficiency of the activities in various fields of functionality. The capability to assess the current level of performance indicators and predict their trend is an important factor in determining the university's development policy and as result, the ranking level of the institution in the world.

To successfully promote universities and research institutions in the international educational and scientific community and increase the interest of foreign researchers, professors, and students, it is necessary to clearly understand the current level of internationalization and advantages that make them attractive for research and students, and weaknesses, which do not allow competing with leading educational and scientific institutions.

The adequate and reasonable use of metrics allows assessing the scientific progress of both individual researchers and institutions are an actual task for analytics. This is especially important when it is necessary to take into account the level of international activities when making decisions about investing in science by the state budget or international and national funds.

Currently, there are a large set of criteria and different methods of using these criteria for evaluation "international", "innovative", "educational", "scientific" and other activities of institutions. These sets of criteria cover all aspects of operations of the educational and scientific institution, namely: administrative, educational, research, innovation, foreign economic, upbringing. Moreover, the quantity of criteria included in each group depends on the evaluation methodology that will be used.

It should be noted that there is no unified list of criteria that would univocally characterize a listed above group. This is clearly shown in the evaluation of research activities by both the scientists and research teams [1].

Most evaluation approaches are based on questionnaires of universities and research institutions. It is a quite common method when ranking universities and research institutions. However, this approach has some limitations mainly related to the need to fill out the questionnaire manually. Currently, there are a large number of open sources available for analysis, such as repositories of publications and databases with information on the economic activity of the institutions. Taking into account the specific environmental requirements, the complexity of the problems of the system integration has determined the necessity to develop methods and tools for designing a distributed information system for analytical tasks, which is based on distributed software architecture for analytics scenarios [2].

Evaluation of information on international activities

The concept of "international activity" of universities and research institutes is a rather "vague" criterion and, at the moment, has no clear definition. Experts note that depending on what purposes pose a university or research institute in this analysis, this will depend on the number of criteria used for evaluation, and the method of calculating the composite indicator.

It should be noted that the criteria for evaluating the international activity of universities and research institutions often differ, although they have much in common. These characteristics, in the form of digital indicators, are presented in many methods for international activities assessing.

However, the main indicators of the international activity of an individual scientist and institution remain: the total number of publications; total number of citations; journal impact factor; maximum citation of one work; Hirsch index.

These indicators have several disadvantages. As noted by the authors [3], the citation index does not show what part of the publication was cited, as well as the author's contribution to this publication. Given the variety of factors influencing the level of citation, it is obvious that the citation index alone cannot serve as a criterion for assessing the quality of the scientific publication.

That is why the problem of choosing criteria for evaluating international activities, developing methods for searching and extracting numerical values of criteria from available sources of information, developing methods for assessing a comprehensive evaluation criterion based on the contribution of each criterion remains relevant and in demand.

When an information-analytical system for analysing the level of international activity developing, it is necessary to expand the approach to assessing the international scientific activity of engineering, scientific and educational institutions and structure those indicators which, to some extent, determine the final value of the indicator. This is especially important in cases where, based on these assessments, decisions are made about the prospects for continued funding of certain areas in research or educational institutions, the need to expand scientific cooperation in a particular scientific field.

The variety of data types in scientific and technical international cooperation is significant - up to several hundred [4]. This raises the dilemma of choosing a basic set: a broad list facilitates a more detailed assessment of individual participants in international activities but makes it difficult to compare with other subjects of international scientific and technical cooperation, as most data values, and the indicators evaluated, may be uncertain due to unavailability of data. Therefore, the choice of the basic set should be made taking into account the intended purpose of the indicator, the availability of data, and the technical capabilities for the extraction and processing of such information.

It is necessary to consider all the information needed to assess the level of international activity as structured information, the essence of which reflects the most important elements of the analysis of the level of international activity, to further determine the ratio of these assessments for each participant.

Scenarios for collecting and analyzing information

There are many different definitions of "scenario". Some researchers understand the scenario as a static picture of the future, others – as a sequence of events in a timeline; also highlighting many differences in other characteristics of the scenario.

At present paper we will accept the most general definition of the scenario: the scenario is the consecutive description of alternative possible variants of events in the future [5].

The successful application of the scenario approach by several different companies, as well as its growing popularity in public strategic management, has given it considerable popularity and prevalence. One of the most important features of the scenario approach is its versatility, flexibility, the ability to apply a variety of methods at different stages of scenario research and process management. The development of a scenario approach using some analytical methods will increase the reliability and validity of the results. This increases the efficiency of its application in the strategic management of complex information and analytical systems.

Scenario can be used to analyze a complicated information environment in which there are many meaningful, interconnected and alternative sources of information. In this case, the scenario itself may include a set of possible events which determine the development of certain factors that affect the outcome of the activity.

The development of system for the analysis of international activities of institutions under the typical scenario for the information retrieval system at the request of the analyst can be understood as a sequence of actions (steps) of the information retrieval system selected by the analyst to solve this problem. Typical scenarios can be changed and transformed while searching for information, followed by saving new versions and versions of versions. After conducting analytical research, new copies of versions can be stored and become separate versions of standard scenarios for further use by the analyst to solve similar analytical problems.

One of the main problems of the system for analysis of international activities during the collection and processing of primary information is its incompleteness and improper relevance in the vast majority of primary information sources. This leads to the need for duplication in determining the sources of information, analysis of information quality during information retrieval and the need to build complicated algorithms for the sequence of analysis and processing of information based on typical scenarios of information analysis [6]. The complexity of such procedures requires not only the adaptation of the structure of typical scenarios to the specific task set by the analyst, but also the adaptation of the parameters of individual elements that implement certain steps of the scenario to these features, both during analysis and after its completion.

The structure of the software implementation of the system for analysis of international activities

To build an information-analytical system for assessing the level of international activity of individual scientists, groups and institutions, built on any criteria, we suggest the system structure which is based on the principles of scenario approach using adaptive microservice architecture with machine learning elements. Schematic diagram of the implementation of a system is presented in Figure 1.

Micro service architecture is a set of independently deployed services. The system constructed in this way can evolve in parts, primarily because each microservice is largely autonomous. In addition, this architecture allows flexible scaling of information system components, thereby ensuring optimal use of existing server equipment. The scaling of each micro service is performed independently. The computing resources are allocated only on demand for a particular functionality of the system and the real user activity.

Micro service architecture is an approach to developing an application that involves abandoning a single, monolithic structure. That is, instead of executing all the limited contexts of the application on the server through internal process interactions, we use several small applications, each of which corresponds to some limited context. Moreover, these programs can run on different servers and interact with each other over the network, for example via HTTP [7].

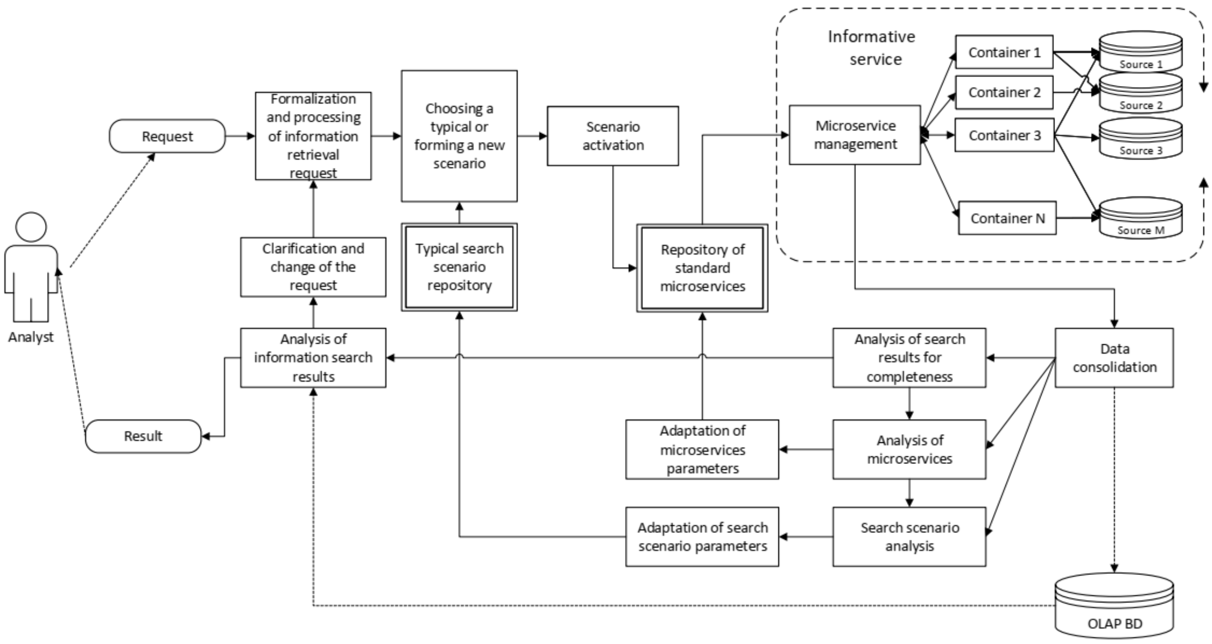

Fig. 1. Schematic diagram of the information-analytical system for assessing the level of international activity

One of the important advantages of micro services is that we can use different technologies to solve the same problem. For example, use different parsing libraries or different data storage or processing tools in each micro service. This makes it possible to build adaptive to the type of task performed software that implements certain elements of machine learning. Some micro services and libraries can lose control in the operation [8]. So, choose a basic set of tools and turn to others only when you need it.

In addition, this approach makes it possible to build and use patterns of extraction and primary processing of information as part of the scenarios for search and processing the information in specific tasks of assessing the level of international activity.

The general process of the information-analytical system for assessing the level of international activity can be divided into the following main steps:

1. Query processing and selection or developing the scenario.
2. Formation the structure of microservices to process the inquiry.
3. Collection and processing of information upon request.
4. Consolidation and preservation of results for further analysis.
5. Analysis of microservices and adaptation of their parameters.
6. Evaluation of the scenario and its adaptation for specific analysis tasks.
7. Analysis of the received information and results of the request.

The first step is the analysis of the request, if necessary, its clarification and formalization of the request in terms of its feasibility in the information and analytical capabilities of the system functions. For a refined and formalized query, a typical scenario is selected from the standard scenario repository, and, if necessary, a new scenario is formed, usually based on one of the standard scenarios.

In the second step, the appropriate microservices are connected to the stage as elements of performing the functions of collecting and processing information according to a typical scenario or its version. This scenario activation procedure is performed using a standard microservice repository.

Thus microservices are developed as separate containers. The software platform Docker is used for rapid development, testing and deployment of microservices [9]. Docker packs the software into standardized blocks called containers. Each container includes everything needed to run the program: libraries, system tools, code and runtime. Docker can quickly deploy and scale microservices in any environment.

The third step is the collection and processing of information on request. This uses the Kubernetes platform, which allows to implement and use applications in microservice architecture, creating a level of abstraction on top of a group of hosts [10]. The Kubernetes platform allows to:

1. Deploy containers and all operations to run the required configuration. These includes restarting stopped containers, moving them to allocate resources to new containers, and other operations.
2. Scaling and launching multiple containers simultaneously on a large number of hosts.
3. Balancing multiple containers during startup. To do this, Kubernetes uses an API to logically group containers. This makes it possible to identify their pools, set the placement, as well as evenly distribute the load.

The fourth step involves the consolidation of information obtained from various sources. Consolidation is a set of methods and procedures aimed at combining data from different sources, ensuring the required level of quality, converting to a single format in which it can be uploaded to the data warehouse [11].

To save the received information after its consolidation, the use of OLAP database is provided, which gives a number of advantages for further processing depending on the need of the analyst who provided the request [12].

The fifth step determines the procedures for analyzing the work of microservices and adapting their parameters. Adaptation of microservice parameters is actually the correction of parameters for standard microservices from the repository of microservices that leads to emergence of new versions of these microservices which will be fixed on standard scenarios as a result of which those versions appeared.

The sixth step is to assess the performance of the scenario and its adaptation to the specific tasks for the analysis of the level of international activity. To build new versions and copies of typical scenarios, it is proposed to use an algorithm based on a linear stochastic automaton [11]. Structural adaptation based on the use of linear stochastic automata is based on the assessment of the most efficient sources of information for a particular request and methods of information processing. This allows to choose the sources of information correctly according to the request and methods of processing this information.

– Two groups of methods are the most critical for the system of assessing the level of international activity:

– alternative methods of information retrieval that are directly related to the relevant sources of information;

– alternative processing methods (often parsing).

For the first group, an important parameter of the quality of work that can affect the adaptation of the scenario is the number of relevant results obtained over time.

For the second, the average processing time per unit of information according to a specific query.

The seventh step concerns the analysis of the received information and results of inquiry. Depending on the assessment of the quality of the received information and its compliance with the request, the analyst, in case of his dissatisfaction, clarifies or changes the request, and otherwise the processing of the request is completed.

Conclusions. The article considers the general structure of the adaptive system of request processing, as developed in the framework of the information-analytical system for assessing the level of international activity. The construction of the query processing system is based on the use of microservice architecture. The placement of microservices is implemented on a distributed network. The extensive system of hosts of the network on which microservices are located is focused on the possibility of more efficient extraction of information from a specific source from which the microservice receives information. The system provides microservices as stand-alone containers. Each of the containers is designed to perform certain operations to extract the necessary information from a specific information source with parameters that can be adapted to a specific request. For the formation and further management of containers, the use of Docker and Kubernetes platforms is envisaged, which allows, in addition, to control the load on each host.

Each request after its verification for feasibility and formalization is implemented in the form of a scenario based on the use of a repository of standard scenarios. The subsystem provides the ability to adapt such scenarios due to the availability of alternative microservices which can be included into the scenario.

The adaptive query processing system presented in the article, which is developed in the framework of creating an information-analytical system for assessing the level of international activity, can be used in the development of various information-analytical systems focused on collecting information on a distributed network in real time.

References

1. Михайлов О.В. Критерии объективной оценки качества научной деятельности // Вестник российской Академии наук. ISSN: 0869-5873. 2011. Том 81. №7. С. 622-625.

2. Kuzminykh V., Koval, O., Voronko M. Standard Analytic Activity Scenarios Optimization based on Subject Area Analysis // CEUR Workshop Proceedings. ISSN: 1613-0073. 2019. Vol. 2577. P. 37-46.
3. Garfild E. «Science citation index» – a new dimension in indexing. Science. 1964. Vol. 144. P. 649-654.
4. Lane J. Let's make science metrics more scientific. Nature, 2010. №464. P. 488-489.
5. Van Notten Ph. Scenario development: a typology of approaches // Think Scenario. Rethink Education. OECD, 2006. P. 69-84.
6. Koval O., Kuzminykh V., Otrokh S., Kravchenko V. Optimization of Scenarios for Collecting Information Streaming Wide-Area Network // 3rd International Conference on Advanced Information and Communications Technologies (AICT), Lviv, Ukraine, 2019. P. 213-215.
7. Ford Neal, Parsons Rebecca, Kua Patrick. Building Evolutionary Architectures: Support Constant Change. O'Reilly Media, 2017. 332 p.
8. Ричардсон Крис Микросервисы. Паттерны разработки и рефакторинга. СПб.: Питер, 2019. 544 с.: ил. ISBN 978-5-4461-0996-8
9. Pethuru Raj; Jeeva S. Chelladhurai; Vinod Singh. Learning Docker. Packt Publishing, 2015. 240 p. ISBN 978-1-78439-793-7
10. Лукша М. Kubernetes в действии = Kubernetes in Action. М.: ДМК Пресс, 2018. 672 с.
11. Kuzminykh Valeriy O., Koval Alexander V., Osypenko Mark V. Methods of Machine Training on the Basis of Stochastic Automatic Devices in the Tasks of Consolidation of Data from Unsealed Sources // CEUR Workshop Proceedings. ISSN: 1613-0073. Vol. 2067 urn:nbn:de: 0074-2067-8. 2017. P. 63-68.
12. Codd Edgar F. Providing OLAP to User-Analysts: An IT Mandate // Computerworld. T. 27. № 30.