AGENT TECHNOLOGY SITUATIONAL EXPRESS ANALYSIS IN ASSESSMENT OF TECHNOLOGICAL DEVELOPMENT LEVEL OF THE BRICS COUNTRIES

D.I. Koshlan[1,2,a], E.S. Tretyakov[3,b], V.V. Korenkov[1,4], B.N. Onykij[3], A.A. Artamonov[3]

[1] *Joint Institute for Nuclear Research, 6 Joliot-Curie, Dubna, Moscow region, 141980, Russia*

[2] *Dubna State University, 19 Universitetskaya, Dubna, Moscow region, 141982, Russia*

[3] *National Research Nuclear University "MEPhI", 31 Kashirskoe shosse, Moscow, 115409, Russia*

[4] *Plekhanov Russian University of Economics, 36 Stremyanny per., Moscow, 117997, Russia*

E-mail: [a] dkoshlan@jinr.ru, [b] etretyakov@kaf65.ru

Stages of development and operation of specialized agent system concerning collection and analysis of the BRICS countries' scientific publications are considered in this paper. The data are extracted from more than sixty sources of authoritative publications in the fields of Chemistry, Physics, Genetics, Biochemistry, Geology etc. Algorithms for data analysis used in the system directed to reveal scientometric indicators and factographic information. The fact analyzed scientific publications are indexed by a referential database Web of Science indicates credibility level of the material. Aggregation of the material is done in a centralized database. The work result let to assess the development level of certain technologies and research in the BRICS countries in a short time. And since the system has a certain degree of autonomy a constant monitoring of scientific and technical activities in the BRICS countries is possible. It is concluded that the use of agent technologies in this field significantly accelerates the analysis of scientific and technical publications in comparison with manual mode.

Keywords: agent system, retrieval system, information and analytical system, situational analysis

1. Introduction

Information products in the modern world are especially in demand among both state and commercial organizations that conduct research and scientific activities. In practice, information products with an analytical overview with macroeconomic estimates based on a large amount of data obtained from various sources are especially valued. Such information products, usually, have a short "shelf life", i.e. a short period of relevance, what means the preparation process should be carried out in a very short time.

This article discusses the method of "situational rapid analysis" using agent technology. The word "analysis" in this case, in contrast to the classical interpretation by ancient Greek philosophers, has the meaning of research in a broad sense, i.e. includes both the method of studying the subject by its constituent parts (analysis), and the study of the subject in the aggregate (synthesis). Agent-based technologies are widely used to conduct such research as a tool for collecting information.

An agent (in this article) is a program with a sufficient level of autonomy to carry out information gathering activities from pre-designated resources and a specific launch schedule.

Currently agent technologies are widely used in such areas as information retrieval systems, news aggregators, ticket sales centers [1]. With the help of agent-based technologies, information retrieval work can be carried out in the Global Network automatically without human intervention, what significantly reduces the time for searching and processing information [2-3]. This fact reveals the meaning of the term "express" in the method.

Situational analysis is an informational and analytical study describing the state of some large system at given time. The task of the situational analysis is to give answers to the questions posed for a specific time period. Situational analysis is conducted to assess the current situation in different areas and tasks [4-5].

The article describes how to solve problems of a particular class. This class of tasks concerning the study of scientific and technical progress level of the states and based on the study of their publication activity in specified thematic areas. An analysis of the scientific and technical level in the BRICS countries in the following areas: physics, genetics, biochemistry, chemistry, geology, astronomy, etc. is cited as an example.

2. Methodology

A preliminary stage of software development is to study the structure of information resources for identifying objects of interest and drawing up a future data model and includes three stages: Reconnaissance, Scanning and Statement.

The task of Reconnaissance is the identification of goals and the study of the legal field. As a result of the analysis of a typical publication page, the following objects of interest were identified: article title, list of authors, organization, key research areas, publication date, abstract. All information sources used for the task allow robots to access content on the http server.

The Scan task is to build an algorithm for collecting target information. For this task, more than sixty agents are created equal to the number of journal-resources. The main link leads to a section of the journal site containing a list of publications of interest. To extract objects of interest, the program automatically follows the link of each publication and collects its attributes.

The task of the Statement is to test individual parts of the algorithm in practice, to identify and list the dependencies. The result of studying an information resource at the Stage of Statement is to obtain a complete list of used software modules and testing of individual parts of the algorithm where these modules are involved.

The further work of creating software for the automated collection of information from various Internet resources is to create an algorithm for the agent's actions. The software development process involves three steps:

1) Creating of the main database tables.

"Agents" and "Publications" tables are created in the database at this stage. The "Agents" table contains a cluster of information resources of the subject area from which relevant information is extracted and indicates the path to the attributes of publications (title, author, date of publication, brief

description and content of the article, scientific area, etc.) for each site. The “Publications” table contains fields for storing retrievable attributes.

2) Development of the main control loop of the process for collecting and processing information.

At this stage, the program logic is implemented and modules for the program operation are developed according to the following structure: scanning the list of sites from the “Agents” table, extracting the attributes of publications and storing them in the “Publications” table of the database to be able to continue working with them.

3) Formation of the resulting Excel-tables.

The result of the system is formed at this stage. With the help of sql-queries the resulting excel-tables are formed that allow to carry out a full-fledged analysis of scientific activity in the BRICS countries.

The algorithm for developing software for site monitoring is presented in the Figure 1.

Thus, after starting the software, the system accesses the database to identify agents that are ready for the session. Then the system collects information for each agent, after what the last articles are saved with a check for uniqueness by link to the "Articles" table of the database. All statistics on downloaded publications can be uploaded to excel files.

3. Results

As a result of the conducted research, the following objects has developed:

1. The main control loop for data collecting and processing, containing the algorithm of the search agents.

2. An integrated database on publications of scientists in the BRICS countries (including scripts for creating a database and corresponding data views in Excel format).

3. Set of final excel-tables.

As a result of the software work for monitoring sites, the information sources of subject area were scanned and the data on the extracted publications were saved in the appropriate database table. The table contains data on 30196 articles for 2016 and 2017 in total.

Agent technologies were applied to solve the problem of forming a specialized system for assessing the scientific activities of the BRICS countries. Also, the algorithm of the system was developed; search agents were configured for a list of information sources; an outline for interacting with the database for collecting, storing and processing information was developed; the resulting Excel tables containing information for assessing scientific activity of the BRICS countries were obtained.

Also, the data from the Web of Science on the publication activity were processed by using internal agents for data analysis. Figures 2 and 3 show fragments of tables with obtained statistics.

4. Conclusion

Created software for monitoring sites to obtain information of interest greatly accelerates and simplifies the analyzing process of scientific activities in the BRICS countries. And since the system has a certain degree of autonomy, constant monitoring of scientific and technical activities in the BRICS countries is possible.

The formed specialized system concerning assessment for scientific activity of the BRICS countries with the use of agent technologies allows to determine thematic areas of the BRICS countries with the most intensive publication activity, to find scientific organizations with the largest number of publications, to evaluate scientific work citation of the scientists in a short time.

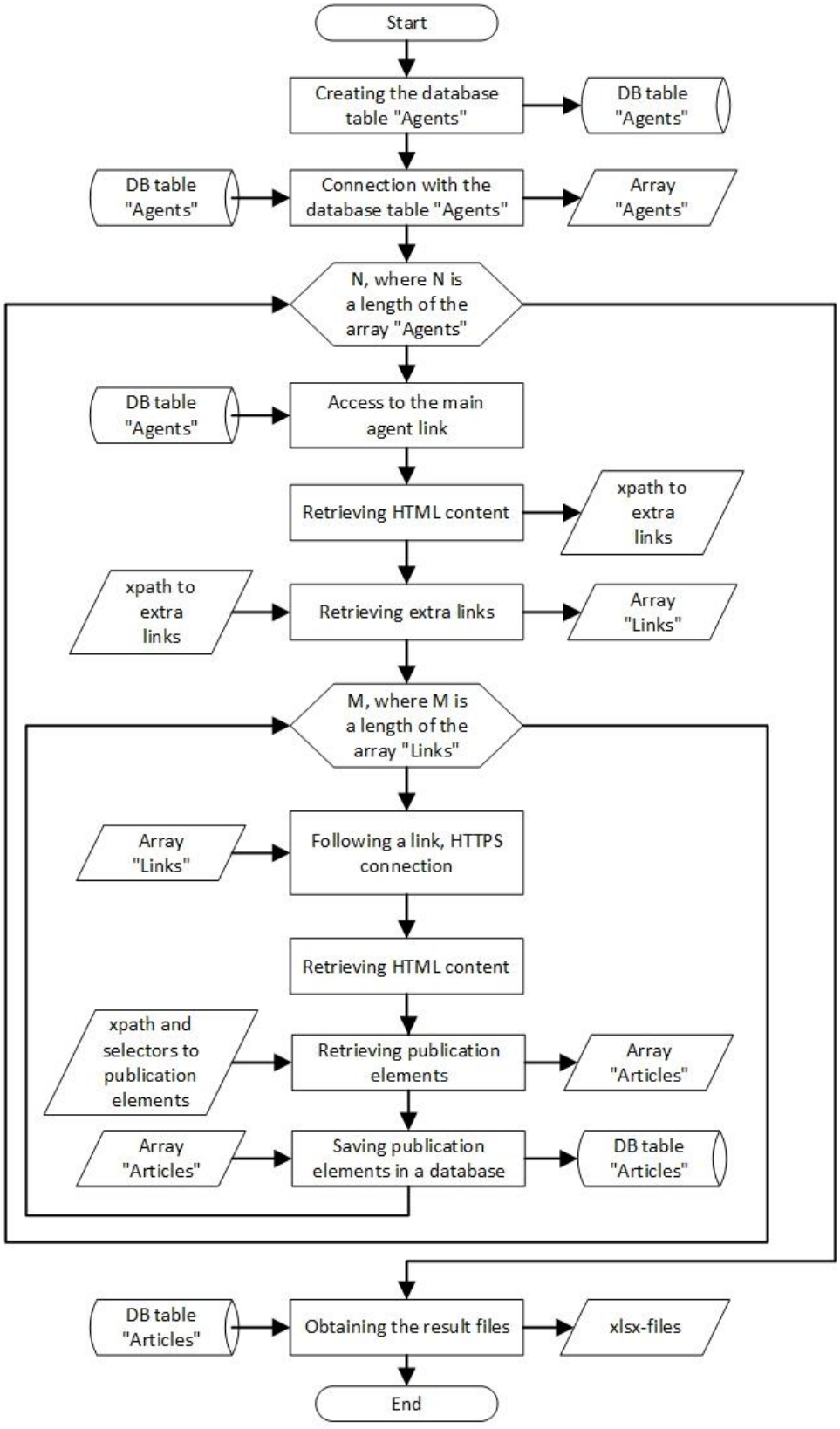

Figure 1. Algorithm for developing software for site monitoring

Affiliation	NumberOf Authors	Number OfArticles	SumOf TimesCited	State
Novosibirsk State Univ, Novosibirsk 630090, Russia	250	4274	26784	Russia
Novosibirsk State Univ, Novosibirsk, Russia	79	3471	26245	Russia
Joint Inst Nucl Res, Dubna, Russia	175	3333	18825	Russia
JINR Dubna, Joint Inst Nucl Res, Dubna, Russia	58	2487	17758	Russia
Natl Res Nucl Univ MEPhI, Moscow, Russia	40	2073	14519	Russia
SB RAS, Budker Inst Nucl Phys, Novosibirsk, Russia	36	1586	13099	Russia
Budker Inst Nucl Phys, SB RAS, Novosibirsk 630090, Russia	76	1737	13032	Russia
RAS, Budker Inst Nucl Phys, SB, Novosibirsk, Russia	27	1356	11303	Russia
Lomonosov Moscow State Univ, Fac Phys, Moscow 119991,	12	44	11264	Russia
RAS, SB, Budker Inst Nucl Phys, Novosibirsk, Russia	18	1301	10573	Russia
Budker Inst Nucl Phys, SB RAS, Novosibirsk, Russia	19	1296	10511	Russia
SB RAS, Inst Nucl Phys, Novosibirsk, Russia	18	1256	10349	Russia
SB RAS, Budker Inst Nucl Phys, Novosibirsk 630090, Russia	50	1337	10056	Russia
Inst Theoret & Expt Phys, Moscow 117259, Russia	82	1339	9124	Russia
Joint Inst Nucl Res Dubna, Dubna, Russia	56	1039	8792	Russia
Moscow MV Lomonosov State Univ, Fac Phys, Moscow 119	47	146	8415	Russia
Budker Inst Nucl Phys, SB, RAS, Novosibirsk, Russia	15	900	7359	Russia
Inst Theoret & Expt Phys, Moscow, Russia	66	1577	7349	Russia

Figure 2. Fragment of the table "Publication activity of scientific organizations"

Year	Science Field	State	Author	Affiliations	Number of articles	TimesCited
2016	Astronomy & Astrophysics	Russia	V., Stolyarov	Univ Cambridge (England); Kavli Inst Cosmo	38	2783
2016	Astronomy & Astrophysics	South Africa	H. C., Chiang	Princeton Univ (NJ 08544 USA); Univ KwaZul	43	2780
2016	Astronomy & Astrophysics	South Africa	N., Bartolo	Univ Padua (Italy); Ist Nazl Fis Nucl (Italy); Af	41	2678
2016	Astronomy & Astrophysics	Russia	R., Sunyaev	Max Planck Inst Astrophys (Germany); Russi	33	2523
2016	Astronomy & Astrophysics	India	S., Mitra	CALTECH (CA USA); Pune Univ Campus (Indi	26	2358
2016	Astronomy & Astrophysics	Russia	I., Novikov	Niels Bohr Inst (Denmark); Niels Bohr Inst (D	33	2327
2016	Physics, Multidisciplinary	Russia	A., Sergeev	Inst Appl Phys (Russia)	6	2229
2016	Physics, Multidisciplinary	Russia	E. A., Khazanov	Inst Appl Phys (Russia)	6	2229
2016	Physics, Multidisciplinary	Russia	O., Palashov	Inst Appl Phys (Russia)	6	2229
2016	Physics, Multidisciplinary	India	S., Bose	Inter Univ (India); Washington State Univ (W	6	2229
2016	Physics, Multidisciplinary	India	V., Boschi	EGO (Italy); CALTECH (CA 91125 USA); Univ I	6	2229
2016	Physics, Multidisciplinary	India	Arunava, Mukherjee	Tata Inst Fundamental Res (India)	6	2229
2016	Physics, Multidisciplinary	India	C., Mishra	Tata Inst Fundamental Res (India)	6	2229
2016	Physics, Multidisciplinary	India	F., Marchesoni	Univ Camerino (Italy); INEN (Italy); Hobart &	6	2229
2016	Physics, Multidisciplinary	India	Archisman, Ghosh	Tata Inst Fundamental Res (India)	6	2229
2016	Physics, Multidisciplinary	India	S., Ascenzi	Univ Roma Tor Vergata (Italy); INFN (Italy); I	6	2229
2016	Physics, Multidisciplinary	India	A., Pai	Indian Inst Technol (India); RRCAT (India); II	6	2229
2016	Physics, Multidisciplinary	India	M., Saleem	IISER TVM (India)	6	2229
2016	Physics, Multidisciplinary	India	A., Samajdar	IISER Kolkata (India)	6	2229

Figure 3. Fragment of the table "Interrelation between the author and the scientific direction"

References

[1] Maleszka M., Mianowska B., Nguyen N.T. (2009) Agent Technology for Information Retrieval in Internet. In: Håkansson A., Nguyen N.T., Hartung R.L., Howlett R.J., Jain L.C. (eds) Agent and Multi-Agent Systems: Technologies and Applications. KES-AMSTA 2009. Lecture Notes in Computer Science, vol 5559. Springer, Berlin, Heidelberg

[2] Onykiy, B., Suslina, A., Ionkina, K., Ananieva, A., Pronicheva, L., Artamonov, A., Tretyakov, E.: Agent technologies for polythematic organizations information-analytical support. Procedia Comput. Sci. 88, 336–340 (2016)

[3] Ananieva, A.G., Artamonov, A.A., Galin, I.U., Tretyakov, E.S., Kshnyakov, D.O.: Algoritmizatiom of search operations in multiagent information-analytical systems. J. Theor. Appl. Inf. Technol. 81(1), 11–17 (2015)

[4] Martin, Wanda et al. "Situational Analysis for Complex Systems: Methodological Development in Public Health Research" AIMS public health vol. 3,1 94-109. 10 Mar. 2016, doi:10.3934/publichealth.2016.1.94

[5] Holloway KA, Kotwani A, Batmanabane G, Puri M, Tisocki K. Antibiotic use in South-East Asia and policies to promote appropriate use: reports from country situational analyses. Br Med J. 2017;358(Suppl 1):9–13.