ATLANTIS PRESS

Advances in Economics, Business and Management Research, volume 90

International Scientific-Practical Conference on Business Cooperation (ISPCBS 2019)

Intellectual Information System Of Subjects Methodological Support Based On The Web Search With Integrated Expert Evaluation

D.V. Grinchenkov
Department «Software computer engineering»
Platov South-Russian State Polytechnic University (NPI)
Novocherkassk, Russia

D.N. Kushchiy
Department «Software computer engineering»
Platov South-Russian State Polytechnic University (NPI)
Novocherkassk, Russia
dkushchiy@rambler.ru

A.V. Kolomiets
Department «Software computer engineering»
Platov South-Russian State Polytechnic University (NPI)
Novocherkassk, Russia

A.N. Shchurov
Department «Software computer engineering»
Platov South-Russian State Polytechnic University (NPI)
Novocherkassk, Russia
dkushchiy@rambler.ru

Thu Thi Nguyen
Sao Do University
Hai Duong Province, Vietnam

***Abstract*—The relevance of the software product of electronic educational resources assessment for the purpose of their further integration into educational process is proved in the article. The basic principles of their search and examination procedure are determined. The general functional structure of intellectual system is considered, basic functions and components assignment of it are described. Architecture of developed information-retrieval system is determined. The basic principles of the procedure of their search are defined. The main screen forms showing work of system of thematic search are given.**

Keywords—information retrieval; web search; subject search; information resources; search algorithm; relevance

I. INTRODUCTION

Tendency of active integration of the electronic educational resources (EER) into educational practice is conditioned by modernization of the higher education under the influence of work market requirements to the prepared personnel and caused by several reasons:

- new scientific and methodical achievements brought a need in a rapid change of content;
- need of an access for students to extensive volumes of data concerning specifics of the studied object;
- need of use multimedia materials, used for better representation of course material.

Used EER shall correspond to scientific and methodical level of the taught disciplines, to requirements of the existing state educational standards, to consider the latest tendencies in science and education. The main advantage of EER is the possibility of their use in all forms of education, including the teaching with the use of distance learning technologies, as well as the transition from the reproductive learning to creative model, which is characterized by the manifestation of creative abilities in the analysis of simulated situations and developing solutions for tasks [1].

II. METHODS

Web-based search of educational content is a special case of a subject search problem. He demands in addition adjusted tools for achievement of optimum results. As example it is possible to consider two interconnected tasks: search of digital educational resources [2-4] to texts of working programs and the automated formation of working programs of discipline on the basis of resources of network. In both cases on linguistic methods need of processing of partially structured text is imposed [5-7].

Let $IR = \{ir_1, ir_2, \ldots, ir_l\}$ – set of Internet resources, $RR = \{rr_1, rr_2, \ldots, rr_m\}$ – set of resources relevant to request, $ER = \{er_1, er_2, \ldots, er_n\}$ – set of resources, approved by experts. The interrelation of the specified sets can be presented as follows:

$$ER \subset RR \subset IR \Leftrightarrow \forall x \left(x \in ER \Rightarrow x \in RR \Rightarrow x \in IR \right).$$

One of key tasks of IR initial set analysis is definition of a way of formation of its subsets RR and ER. The general idea of creation of a set ER consists in a combination of formal methods of search and information processing and statistical methods of formation of the integrated assessment based on heuristic knowledge of experts.

For realization that approach, it is rational to develop the intellectual information system (IIS) [8] which modular and functional structure is presented in Fig. 1.

Above-mentioned activities to be implemented following tasks should be solved:

1. Development of complex assessment method of expert's opinion with use of the integrated indicators

allowing creating single information and analytical model of decision support process.

2. Development of intellectual system as instrument of monitoring, analysis and forecasting of indicators in the form of software product.

3. Choice of structure and data model for storage of subject domain knowledge.

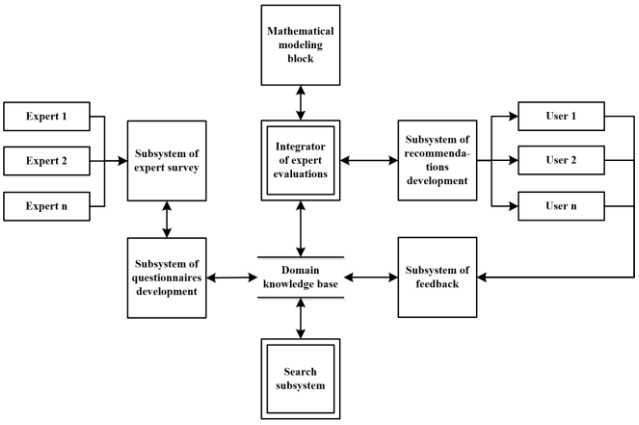

Fig.1 Modular and functional structure of a methodical providing formation subsystem.

4. Definition of method for questionnaires development, for results processing and structuring.

5. Development of procedures connected with involvement of experts and use of their knowledge.

6. Development of a specialized search subsystem for domain knowledge base.

Effective work of a subsystem for the integrated assessment of expert's opinion development that is built in decision-making system could be realized by means of Delphi method combination and statistical data processing.

The main components of system have the following assignments:

1. Integrator of expert evaluations – system kernel providing interaction between domain knowledge base and mathematical modeling block for users decisions support.

2. Mathematical modeling block of integrated expert evaluations realizes is for Delphi method realization.

3. Subsystem of expert survey organize work with remoted experts, as well as realizes the survey interface.

4. Subsystem of questionnaires development provide with systematization of primary information based on statistical analysis and data transformation to the formalized forms with the subsequent formation domain knowledge base.

5. Subsystem of feedback with users realizes the interface specific user's interactions with domain knowledge base.

6. Subsystem of recommendations development for optimal decision-making is intended to interpret results of user inquiries processing.

7. Search subsystem of domain knowledge base formation.

IIS database is storage of the resources expecting expert assessment. It is created by means of rather independent operation of a retrieval subsystem.

The main difference of analytical search engines as offered system from simple search engines is in the query view representation and ranging of the found resources taking into account a context of the chosen subject domain [9-11].

Developed solution performs the following basic functions as any other retrieval systems [12] (Fig. 2):

1. Indexing – collection of electronic resources and creation of their logical images for internal representation in system, with the subsequent storage of the received images

2. Query view – descriptions of user information needs in the language supported by a search engine [13].

3. Comparison – calculations of closeness estimate (relevance) between inquiries and documents. The set of results is formed on the basis of closeness estimate which then comes back for creation of the integrated expert assessment (block 4 Fig. 2) and obtaining pertinent list of electronic educational.

For creation of mathematical model of parametrical optimization of inquiry it is offered to use mathematical model of the information and analytical search engine:

$$Model = [Q', U_{IR}, R], U_{IR} \to R ,$$

where Q – set of input values (query to the system); R – set of output values (search results); $IR = \{IR_i\}$ – set of Internet resources; U_{IR} – set of all subsets IR; $R = \{R_i\}$ – image of queries set to set of text documents, corresponding the work of information-retrieval machine.

Set of relevant resources found in the result of search forms in the following way:

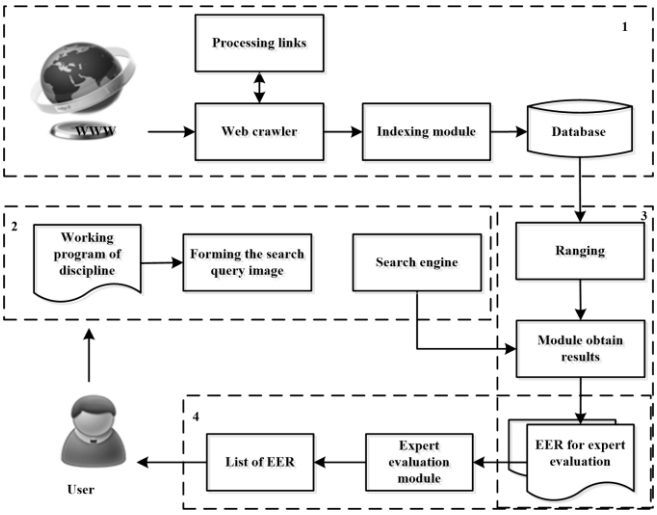

Fig.2 Architecture of developed information-retrieval system.

$$R : \{Q \to U_{IR}, U_{IR} \to R \mid R(q) = IR(q)\} ,$$

The principle of ranging of results of information search is based on the integrated assessment consisting of mechanical relevance of a resource and competence of the expert [14]. Calculation of competence of experts when forming the new educational program it is expedient to use the integrated competence indicator. It is based on the requirements of the Register of the personnel certification system K_1, knowledge of the subject domain of K_2 [15], as well as a posteriori data K_3 – the frequency of occurrence of the proposed keywords and phrases.

For calculation of K_1 the matrix of pair comparisons is filled. Let A_1, A_2..., A_n main characteristics of K_1. For determination of structure of an object the matrix of pair comparisons is filled. If to designate A_i factor share through v_i and v_j (the estimates put forward by the expert in the competence scale), then the matrix element is equal to: $a_{ij} = v_i / v_j$.

In the proposed application of the method of paired comparisons not sizes of differences of values of factors, but their relation are defined. At the same time it is obvious: $a_{ij} = 1/a_{ji}$.

Calculation of values of a competence vector of criterion K_{1i} is carried out on the basis of a formula:

$$AK_{1i} = \lambda_{\max} K_{1i} ,$$

when $\lambda_{\max}$ – the maximum eigenvalue of the matrix A.

Parameter K_2 is calculated using methods of semantic analysis of the list of publications on principles similar to the technology of selection of resources for a thematic search.

To calculate the sub-coefficient of competence of a posteriori data K_3, the following designations are introduced: x_i^t – estimation of the object in the approximation t; x_{ij} – estimation of the object i by the expert j; k_j^t – competence of expert j in approximation t; λ^t – consistency of the estimation of the object in the approximation t.

The process itself has the form of a recurrent procedure:

$$k_j^t = \frac{1}{\lambda^t} \sum_{i=1}^{n} x_{ij} x_i^t ; \ \sum_{i=1}^{n} k_j^t = 1 \left(j = 1, 2 ... m \right) .$$

Calculations begin with t=1. Initial values of coefficients of competence are accepted identical and equal $k_j^0 = 1/m$, where m – the number of experts. Questions of convergence are solved on the basis of the Perron-Frobenius theorem. The general competence of experts in assessing the resources found is calculated by summing K_1 and K_2.

III. Results

Example of work of a subsystem for the formation of methodological support are shown by the screen form given on Fig. 3, 4.

Modification of the available subsystem by creation of graphical representation of semantic network of subject domain of inquiry and realization of the algorithms scaled on large volumes of text collections are planned.

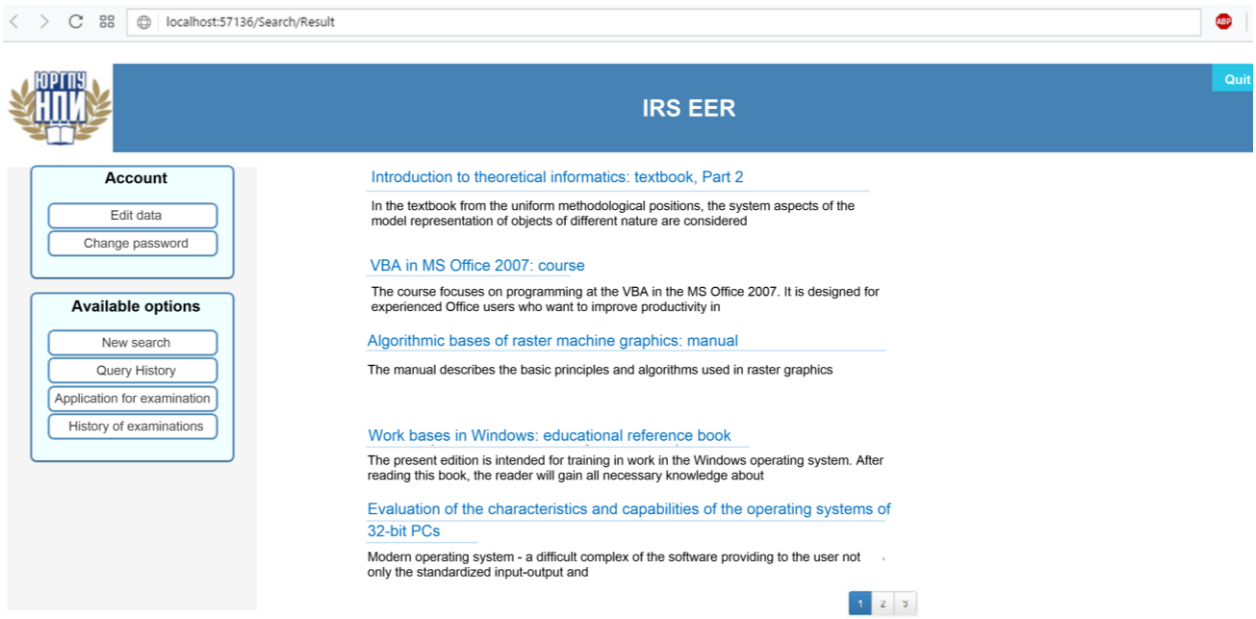

Fig. 3 Search results of electronic educational resources for the set discipline.

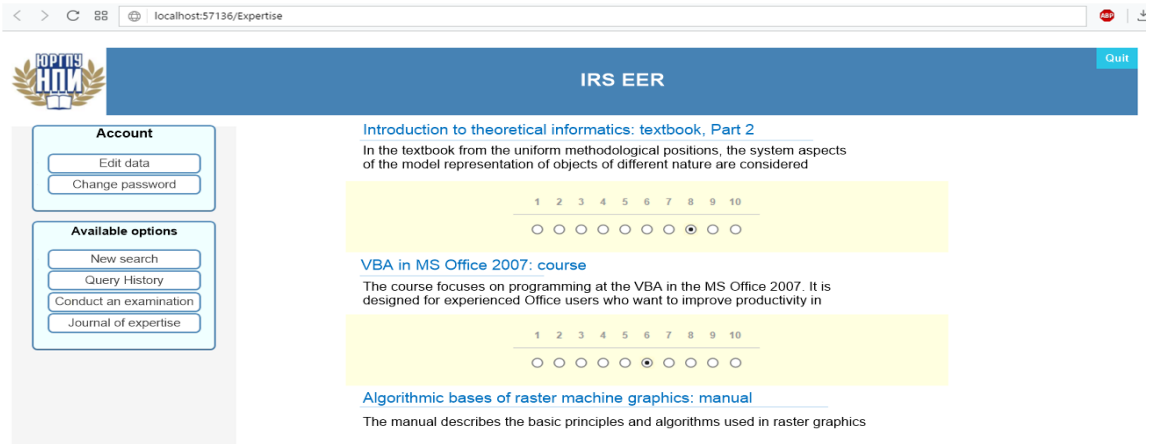

Fig. 4 Assessment of electronic educational resources by the expert.

IV. Conclusion

Comparison of efficiency of simulated system and traditional information retrieval systems has been carried out with the aim of efficiency assessment of the developed models and algorithms. Results were compared to a reference system – the hypothetical system finding all available relevant to this inquiry documents. Comparison was carried out by the number of the relevant documents issued by systems.

It should be noted that distinctions in efficiency are seen in process of increase in volume of the worked out array of data. The collection of pages of Wikipedia (about 2500 documents) was used for simulation. At such size of archive the difference in number of the issued relevant documents makes about 15-20%.

The further plan of work includes working out increase in efficiency of the developed algorithms on large volumes of text collections.

References

[1] D. Grinchenkov, D. Kushchiy, A. Kolomiets, “One Approach to the problem solution of specialized software development for subject search,” Proceedings of the 4th International Conference on Applied Innovations in IT Editors: Eduard Siemens, Bernd Krause, Leonid Mylnikov. 2016, pp. 39-48.

[2] A. Shah, S. Jain, R. Chheda, A. Mashru, “Model for re-ranking agent on hybrid search engine for e-learning,” Proceedings - 2012 IEEE 4th International Conference on Technology for Education, 2012, pp. 247-248.

[3] S. Pudaruth, K. Boodhoo, L. Goolbudun, (2016) An intelligent question answering system for ICT. International Conference on Electrical, Electronics, and Optimization Techniques, ICEEOT 2016, pp. 2895-2899.

[4] M. Chen, M. Dcary, “A Cognitive-Based Semantic Approach to Deep Content Analysis in Search Engines,” Proceedings - 12th IEEE International Conference on Semantic Computing, ICSC 2018, 2018, pp. 131-139.

[5] M. Glava, E. Malakhov, “Searching similar entities in models of various subject domains based on the analysis of their tuples,” 2016 International Conference on Electronics and Information Technology, EIT 2016 - Conference Proceedings. DOI: ICEAIT.2016.7501001.

[6] H.C. Sun, C.J. Jiang, Z.J. Ding, P.W. Wang, M.C. Zhou, “Topic-Oriented Exploratory Search Based on an Indexing Network,” IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 46 (2), 2016, pp. 234-247.

[7] S. Oh, Y.-J. Byon, H. Yeo, “Improvement of Search Strategy with K-Nearest Neighbors Approach for Traffic State Prediction,” IEEE Transactions on Intelligent Transportation Systems, vol. 17 (4), 2016, pp. 1146-1156.

[8] D.V. Grinchenkov, D.N. Kushchiy, A.V. Kolomiets, “One Approach to the Solution of Subject Search Problem of Electronic Educational Resources on the Internet,” 2016 2nd International Conference on Industrial Engineering, Applications and Manufacturing, Proceedings 2. 2016. pages. 7911704

[9] O.A. Nikolaychuk, A.I. Pavlov, A.B. Stolbov, “The software platform architecture for the component-oriented development of knowledge-based systems,” 41st International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2018 - Proceedings, 2018, pp. 1064-1069.

[10] V. Lytvyn, V. Vysotska, D. Dosyn, O. Lozynska, O. Oborska, “Methods of Building Intelligent Decision Support Systems Based on Adaptive Ontology,” Proceedings of the 2018 IEEE 2nd International Conference on Data Stream Mining and Processing, 2018, pp. 145-150.

[11] D.V. Grinchenkov, D.N. Kushchiy, “Architecture of system the subject search of electronic educational resources on the Internet,” J. Informatizaton and communication, no.3, 2016, pp. 143-146.

[12] M. Chen, M. Dcary, “A Cognitive-Based Semantic Approach to Deep Content Analysis in Search Engines,” Proceedings - 12th IEEE International Conference on Semantic Computing, ICSC 2018, 2018-January , pp. 131-139.

[13] X. Cao, Y. Zhang, F. Zhang, C. Ni, “Architecture design of subject-oriented web crawler,” Proceedings - 2013 4th International Conference on Intelligent Systems Design and Engineering Applications, ISDEA 2013, pp. 174-177.

[14] Wei Yang, Zhizhuo Yang, “Query based summarization using topic background knowledge,” 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery, 2017.

[15] Yang L., Hu Z., Long J., “Service of searching and ranking in a semantic-based expert information system. Proceedings - 2010 IEEE Asia-Pacific Services Computing Conference, APSCC 2010, pp. 609-614.