Monitoring the Development of University Scientific Schools in University Knowledge Management

Gulnaz Zhomartkyzy and Tatyana Balova
Department of Information System, D. Serikbayev East Kazakhstan State Technical University, 69 Protozanov A.K., Ust-Kamenogorsk, Kazakhstan

Keywords: Knowledge Management System, Model of a Specialist, Monitoring of Scientific Schools, Clustering, Scientific Community.

Abstract: This paper proposes a technological approach to university scientific knowledge management which integrates the ontology based knowledge model and the methods of university scientific resource intellectual processing. The process-oriented On-To-Knowledge methodology is used as the basis for university scientific knowledge management. Some models and methods of university scientific knowledge management have been studied. The developed model of a specialist that reflects the level of scientific activity productivity and overall assessment of the employee's scientific activity has been described. A specialist's competence in knowledge areas is based on the processing of information resources. The approach to the university scientific school identification based on the clustering of university academic community common interests has been described.

1 INTRODUCTION

A large amount of accumulated information resources and the high speed of new information arrival impose increasingly high requirements to modern systems designed to provide information support to university scientific processes.

The intellectual capital or intangible assets of the university are the source of new scientific knowledge. At the level of organization scientific knowledge or intellectual resources is a complex category which combines intellectual capital, people, and various forms of intangible assets which concentrate knowledge and professional skills (Klimov, 2002).

Russell Akkof, one of the classics of Operational Research, proposed the following hierarchy of knowledge: data - information - knowledge - understanding - wisdom (Ackoff, 1989). There are different approaches to the classification of knowledge in organizations. The most common is the division of knowledge into explicit and tacit knowledge. The transformation of knowledge in an organization occurs through explicit and tacit knowledge interaction. The knowledge conversion or transformation results in its qualitative and quantitative increase.

The notion of knowledge management in an organization is determined by the authors in (Zaim, 2007) as the strategy and the transformation of knowledge.

The main focus in knowledge-intensive organizations is on the creation, transfer and development of knowledge, so effective knowledge management is the matter of survival for such organizations (Miles, 2005; Scarso et al, 2010). There arises a need in processes, infrastructure and organizational procedures at a higher education institution that would allow its employees to use its corporate knowledge base.

This paper considers the problem of monitoring the development of university scientific schools, which is one of functional components of univertity SKMS. The paper discusses some models and methods of university scientific knowledge life cycle support.

Section 2 describes the "University Scientific Knowledge Management ". Section 3 describes "The Monitoring of University Scientific School Development." Section 4 describes "The Development of a University Researcher Model", section 5 presents "The Approach to the University Scientific School Identification", Section 6 is the conclusion.

Zhomartkyzy G. and Balova T..
Monitoring the Development of University Scientific Schools in University Knowledge Management.
DOI: 10.5220/0005464202220230
In *Proceedings of the 17th International Conference on Enterprise Information Systems* (ICEIS-2015), pages 222-230
ISBN: 978-989-758-097-0

2 THE STRUCTURE OF UNIVERSITY SCIENTIFIC KNOWLEDGE MANAGEMENT SYSTEM

In this paper the university's scientific knowledge management system (SKMS) is considered as an aggregate of information, software, technical means, and organizational solutions aimed at efficient management of the university's available intellectual resources and training specialists who meet the modern requirements.

According to the above definition, university scientific knowledge management (SKM) can be understood as:

- the aggregate of processes associated with the creation, distribution, processing and use of university scientific knowledge;
- an established systematic process of working with information resources and knowledge, scientists and specialists in certain areas in order to facilitate access to knowledge and re-use them with modern information technology at the university.

The purpose of university SKMS is the formation of a unique ontology-based integrated intellectual environment to improve the competitiveness of the university's science and education. The university SKMS is the technological component of the university SKM, which provides the creation, organization and dissemination of scientific knowledge among the university staff.

There are following approaches to knowledge management: organizational and technological (Tuzovskiy, 2007). The technological approach puts the application of IT-technologies in line with the organizational measures. The model of technological approach to knowledge management is shown in Figure 1.

The process-oriented On-To-Knowledge methodology is used as the basis for university scientific knowledge management (Sveiby,1989; Staab et al, 2001). The methodology of KMS development and support is based on the process and metaprocess of working with knowledge (KnowledgeMetaProcess and KnowledgeProcess). The basis of the metaprocess of working with knowledge (KnowledgeMetaProcess) is the development of an ontology, which consists of the following steps: a feasibility study, the beginning, clarification, evaluation, support and evolution.

The ontology is the link/(linking element) of knowledge objects and a connecting bridge between different steps of knowledge transformation processes (KnowledgeProcesses). The development of the ontology is the important aspect of knowledge management solution support. The development and deployment of applications of knowledge management takes into account the requirements of "KnowledgeProcess" and considers such processes / issues as:

- metaprocess of working with knowledge (KnowledgeMetaProcess);
- software engineering (software development and design– Software engineering);
- the corporate culture of the organization.

The process of working with knowledge (Knowledge Process) focuses on the use of KM-solutions, i.e. after KM-application are fully realized and implemented in the organization, the cycle of knowledge transformation is performed. The knowledge transformation cycle consists of the following steps: creation, storage, search and access, use.

The developed model of technological approach to knowledge management based on the methodology described above is shown in Figure 1.

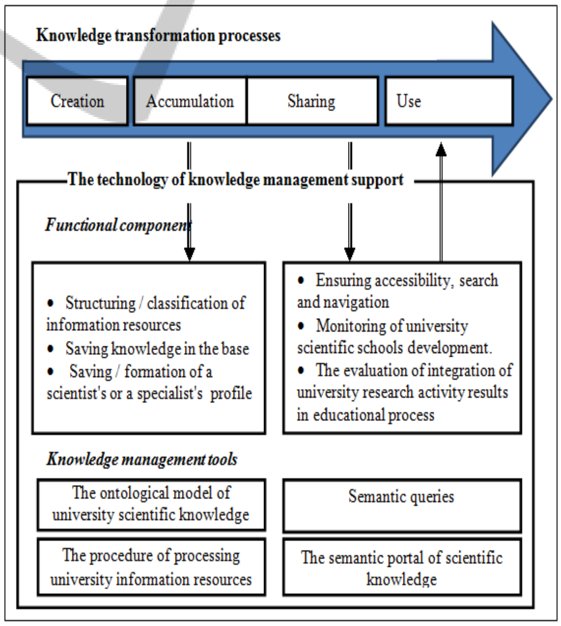


Figure 1: Technological approach to university scientific knowledge management.

The proposed technological approach integrates a functional component and knowledge management tools. The functional component includes:

- classification of information resources,
- a scientist's/a specialist's scientific profile formation and saving;

- identification of university scientific schools and research directions;
- ensuring the availability, search and navigation.

Knowledge management tools are:

- the ontology based university scientific knowledge model,
- the procedure of university information resources processing,
- semantic queries,
- semantic Portal of university scientific knowledge.

The following sections describe the implementation of the task of monitoring scientific school development in university knowledge management.

3 MONITORING THE DEVELOPMENT OF UNIVERSITY SCIENTIFIC SCHOOLS

The priority task of integrating research and education is the development of scientific schools which must be the main result of fundamental science and education interaction. Introduction to the research undertaken by scientists of scientific schools is the best school for young people.

Scientific schools provide constant growth of qualification of their participants; the presence of several generations in bundles of "teacher-student" ensures the continuity of generations (NC STI RK). The development of scientific schools and scientific and pedagogical teams is the basis for the development of fundamental scientific research and training quality improvement of research and educational personnel.

Scientific schools form that dynamic unit of science which ensures the continuity of scientific knowledge and creates optimal conditions for its development. Scientific school is the key element of collective preservation and multiplication of knowledge, one of the conditions to maintain the quality of research, and hence the quality of training scientific personnel. Scientific school is a clearly defined direction of scientific research carried out in the framework of specific scientific specialties (Trubina and Zabelina, 2011).

Identification of scientific schools is becoming increasingly important in recent years in connection with the development of mechanisms for organization effectiveness assessment in tenders for financing projects, their certification and accreditation. The availability of scientific schools is one of the most important criteria for foreign scientific funds which conclude contracts on joint research and grants as well as the criteria taken into account in establishing the rating of organizations.

One of the qualitative characteristics of a particular scientific direction's overall development and potential is the state of scientific schools. The creation, reorganization and coordination of scientific schools are regulated by universities. Monitoring the development of scientific schools remains a major issue in university scientific and innovation activity management. Thus, identification, recording, development, and monitoring the development of scientific schools is one of the priorities of science and education.

We have studied and developed a model of scientific community and the method of its intelligent processing to implement the functions of monitoring. The overall structure of monitoring the development of university scientific schools is presented in Figure 2.

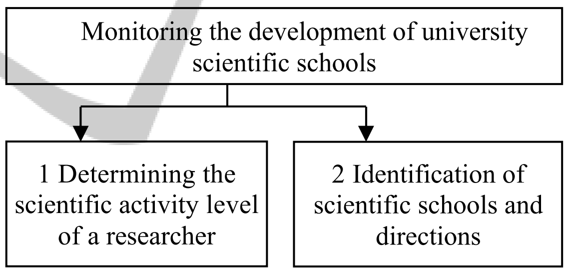


Figure 2: The main functions of monitoring the development of university scientific schools.

The scheme of monitoring shows the following functions: determining the researcher's competence level, identification of scientific schools and research directions.

The following sections describe the model of a specialist which reflects the level of scientific activity productivity based on the calculation of entropy and overall scientific activity evaluation; as well as the approach to the university scientific school identification based on clustering the university scientific community for common interests (university scientific schools).

4 THE DEVELOPMENT OF A UNIVERSITY RESEARCHER MODEL

One approach to human capital management is to develop a model of a university researcher (the model of a specialist).

Currently, there are two ways to create and support the model of a researcher: by a survey (qualification audit in an organization) and by monitoring their work in the knowledge management system (scientific papers, projects) (Tuzovskiy, 2007). The paper supports the definition given in Tuzovskiy , 2007), where the model of a researcher refers to a sound set of interrelated properties of a specialist, which can be formally described and used to support the efficient work with implicit knowledge.

In the scientific knowledge ontology the model of a specialist has the following formal description which includes a set of contextual and content metadata:

$$M_s = \{M_{context}, M_{content}\} \quad (1)$$

where: $M_{context}$ – is contextual metadata of a specialist description; $M_{content}$ – is content metadata, which describe the specialist's competence.

Contextual metadata $M_{context}$ of a specialist include such parameters as:

- identification (a name, a photo, the date of birth, the place of birth, a login, a password);
- contact information (postal and email addresses, a personal web page, phone numbers;
- education (diplomas, certificates, etc.);
- professional achievements (prizes in competitions, awards, medals, etc.).

Content metadata $M_{content}$ provide the description of the specialist's competence as a set of his competence characteristics:

$$M_{content} = \{C_{sa}, C_{cs}, C_{oe}\} \quad (2)$$

where: C_{sa} – the competence of a specialist in fields of knowledge relevant to rubrics which are described as classes in the scientific knowledge ontology O_{SK}.

C_{cs} – a measure of specialist's scientific activity efficiency (the level of the specialist's competence dispersion);

C_{oe} - the overall assessment of scientific work.

The model of an individual researcher's scientific activity is determined by the factors of scientific activity (Figure 3).

In ontological information model these factors are grouped into the following classes: Event, Project, Publication.

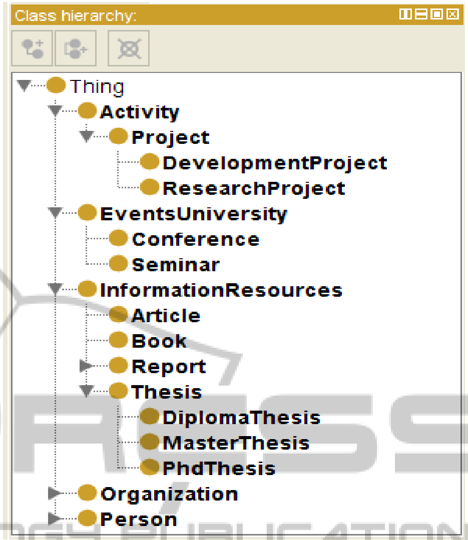


Figure 3: Classes of the information model used to simulate a university researcher's activities.

4.1 A Specialist's Competence in Areas of Knowledge based on the Classifications of Information Resources

Automatic processing of scientific electronic resources by the methods of TekstMining text processing are required to implement the described above model.

A specialist's competence in areas of knowledge C_{sa} is formed on the basis of the specialist's scientific profile. The specialist's scientific profile is based on the classification factors of his scientific activity (publications) by scientific areas (Zhomartkyzy, 2014).

The specialist's scientific profile is determined as the profile of all his publications: the profiles of documents are formed by processing the university scientific resources. The profile of a document is determined as the vector of all its relevant classes:

$$PD = \left(R_{1,\dots,}^{d} R_{c}^{d}\right) \quad (3)$$

where: R_c^d – are relevant classes PD.

Accordingly, the specialist's competence in areas of knowledge C_{sa}, relevant to the specialist's scientific profile is determined as the profile of all his publications.

$$C_{sa} = (PD_1, \ldots, PD_i) \qquad (4)$$

where PD_i are all the author's documents.

The final step of text classification is the formation of the document's semantic profile by creating the class of "Information Resources" individuals of scientific activity ontology.

4.2 The Calculation of the University Specialist's Scientific Activity Efficiency

To assess the university specialist's scientific activities efficiency we suggest using the level of dispersion of his competence.

The model of an individual researcher's scientific activities is determined by the factors of scientific activities. These factors are grouped into ontology classes in the model of ontology.

The connection between the Person class and the class factor of scientific activities is shown below:

$$P_i \equiv P \cap \exists PersonHasIR.IR$$
$$IR_i \equiv IR \cap \forall \text{publHasDivis}.FC_i$$

where, P – is persons, IR – иis information resources, FC_i is the field of knowledge..

Each researcher works in at least one field of knowledge (VINITI rubricator, VINITI - All-Russian Institute of Scientific and Technical Information). Therefore, the classification of scientific activity factors is carried out by means of the VINITI rubricator of fields of knowledge up to level 3.

Cybernetics →Artificial Intelligence → Knowledge engineering

Cybernetics → Artificial Intelligence → Expert systems

Cybernetics→ Theory of modeling → Mathematic modeling

This paper proposes the method for calculating the efficiency index of a specialist's scientific activities C_{cs} to analyze the competence of employees in a particular field. A specialist's scientific activity efficiency C_{cs} is calculated using the entropy.

The more papers of a specialist (researcher) are grouped by a certain category, the lower entropy and the higher the specialist's scientific activity efficiency are (Adamic et al, 2008), (Baesso et al, 2014). A specialist who has a high entropy works, as a rule, in several fields of knowledge, i.e., the specialist has a lower scientific activity efficiency (scientific competence):

$$C_{cs} = -\sum_{i}^{N} P_c \times log_2(P_c) \qquad (5)$$

$$P_c = \frac{P_i}{P} \qquad (6)$$

P_i – is the number of the researcher's papers by heading $i = \overline{1, N}$;

P - the total number of the researcher's papers.

An example of calculations using formula 3 is shown in Table 1.

Table 1: A Specialist's Scientific Activity Efficiency Calculation.

A researcher	Fields of knowledge	The number of scientific resources	The efficiency index of a specialist's scientific activities (C_{cs})
$specialist_1$	Physics of Atom and Molecule	1	1,68
	General Physics	3	
	Solid State Physics (nano-sized objects, the structure of solids, general issues of Solid State Physics)	5	
	Physics of Gases and Liquids	1	
$specialist_2$	Solid State Physics	5	0
$specialist_3$	General Physics	6	0
$specialist_4$	Nuclear Physics	1	0,91
	General Physics	2	
$specialist_5$	General Physics	7	0,54
	Physics of Gases and Liquids	1	

Threshold values of the specialist's scientific activity efficiency C_{cs} were determined empirically:

$0 < C_{cs} < 1$ - a high level;

$1 \leq C_{cs} < 2$ - a medium level;

$C_{cs} \geq 2$ - a low level.

The analysis of results of personal calculation C_{cs} for leading university scientists by knowledge areas "General Physics", "Physics of Solids", "Physics of Atoms and Molecules," "Physics of Gases and Liquids" and "Nuclear Physics" confirms the applicability of the formula for calculating the entropy of the researcher's scientific competence.

4.3 The Calculation of Total Assessment of the University's Specialist's Scientific Activity

The specialist's scientific activity efficiency is a quantitative indicator of knowledge, skills and abilities in a scientific field of a corresponding specialty.

The qualitative analysis, i.e. the definition of a specialist's professional competence, is also needed for making management decisions in the context of the university's different departments.

Each researcher has a trajectory of educational and scientific activities (Figure 4): scientific activities, educational activities, participation in competitions and grant projects, international mobility.

This scheme allows us to calculate the qualitative characteristics of the overall assessment of the specialist's scientific activity.

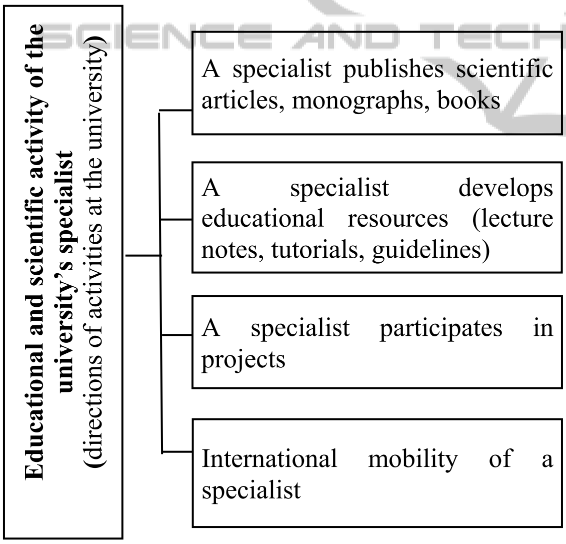


Figure 4: Directions of scientific activities at the university.

The calculation of a specialist's overall scientific activity evaluation $- C_{oe}$ can be shown schematically as follows (Figure 5).

A university specialist with a high overall assessment of scientific activity:

- Works in all areas of scientific activities;
- Has a high index of scientific competence C_{cs}.

A university specialist with a medium overall assessment of scientific activity:

- Only develops educational courses or is only involved in projects;
- Has a medium index of scientific competence C_{cs}.

A university specialist with a low overall assessment of scientific activity:

- Only develops educational courses;
- Has a low index of scientific competence C_{cs}.

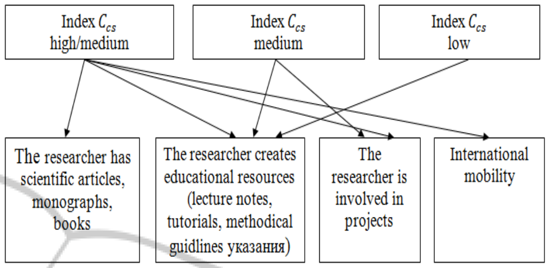


Figure 5: The scheme of the calculation of a specialist's overall scientific activity evaluation.

5 THE APPROACH TO THE UNIVERSITY SCIENTIFIC SCHOOL IDENTIFICATION BASED ON THE UNIVERSITY COMMUNITY CLUSTERIZATION BY COMMON INTERESTS

Each scientific school or scientific direction forms a scientific community by interests and develops in accordance with some specific rubrics of knowledge areas (Cantador and Castells, 2011). The VINITI rubricator of knowledge areas is used as the rubricator. In the proposed approach, a *model of the scientific community* is described as follows:

$$M_{SC} = \{r_{1.1}, \quad r_{1.2}, \quad r_{2.3},\} \tag{7}$$

where r_i are rubrics corresponding to specific areas of science and technology (one subrubric may be in several scientific fields).

The model of a scientific community is shown in Figure 6.

The proposed approach requires to carry out the university *scientific community* clustering based on its memebers' common interests to identify scientific schools and research directions.

To identify scientific schools and research directions DBSCAN clustering method is used in the scientific community model.

The principal advantages of this method served as the basis of the choice of DBSCAN density clustering method:

- Identification of the number of clusters (based on the notion of point density);
- The clustering algorithm is able to detect clusters of different shapes;
- Resistance to noise objects.

The idea underlying the algorithm is that within each cluster there is a typical dot density (of objects), which is significantly higher than the density outside the cluster (Figure 7). The density in the areas with noise is lower than the density of any of the clusters. For each dot of the cluster its neighborhood of a given radius must be at least a certain number of points, this number of dots is specified by a threshold value. (Bolshakova et al, 2011), (Marmanis et al, 2011).

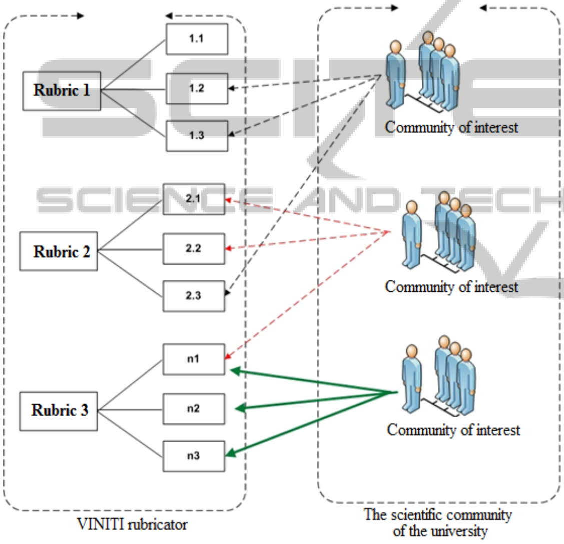


Figure 6: A university scientific community by interests.

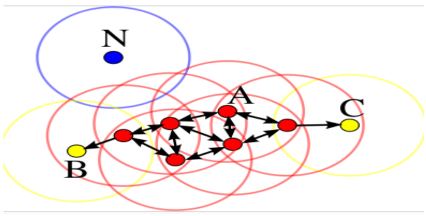


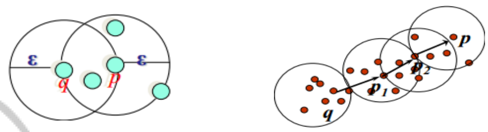
Figure 7: Example of a cluster of arbitrary shape.

In Fig.7 A is a core point. B, C are border points.Cluster C_j is not the empty subset of objects satisfying the following conditions, at given *Eps* and $MinPt$, where Eps is the maximum distance between adjacent points, MinPt is the minimum number of neighboring points:

- $\forall p, q$: if $p \epsilon C_j$ *and* q is density-connected from p, then $q \epsilon C_j$, at given *Eps* and *MinPt*;
- $\forall p, q \epsilon C_j: p$ is density − connected with q, at given *Eps* и *MinPt*.

Figure 8 is given below for detailed description of the points.

Thus, a cluster is a set of closely-related points. Each cluster contains at least *MinPt* of documents. To perform clustering the model of a scientific community is translated into a binary matrix (Figure 9). The values of matrix elements correspond to the presence or absence of work on the appropriate rubric.



a) density-reachable points

b) density-reachable points and density-connected points of the class

Figure 8: Types of points that form classes in DBSCAN algorithm.

Members of the scientific community	Rubrics								
	0	0	0	0	0	1	0	0	1
	1	0	0	0	0	1	0	0	1
	1	0	0	0	0	1	0	0	1
	1	0	1	0	0	0	1	0	1
	1	0	0	0	0	1	0	1	1
	0	0	0	0	0	1	0	0	1
	1	0	0	0	0	1	0	0	1

Figure 9: The matrix of the scientific community model description.

The clustering algorithm based on the density of points is described below (Bolshakova et al, 2011).

Input: a set of objects Q, parameters - Eps (the distance between the objects of the class), $MinPt$.

1. Determination of directly density-reachable points:

$$p \in N_{eps}(q), \ |N_{eps}(p)| \geq MinPt$$

where, q is a core point, p is a border point. Point p is directly density-reachable from point q,

2. Determination of all density-reachable, density-connected points of the current class:

$$p \leftarrow p_{(i+1)} \leftarrow p_i \leftarrow q$$

$$p \in N_{eps}(p_k), \ |N_{eps}(p_k)| \geq MinPt$$

Output: a set of clusters

Noise is a subset of objects that do not belong to any cluster,

$$p \in Q | \forall j \notin C, j = 1, |C|.$$

The value of *Eps* is determined as the distance between the researchers' profiles. A researcher's profile consists of all its relevant rubrics and is presented as a vector. VINITI rubricator of knowledge areas is used as a rubricator. The set of vectors forms a matrix of researchers's profiles . To calculate the distance a cosine measure of adjacency is used. The value *MinPt* is the minimum number of the subjects of scientific school, i.e. the subjects of "the communities of interest" in the model of a scientific community.

For approbation of the proposed approach we chose scientific communities of D. Serikbayev EKSTU and Ioffe Physical-Technical Institute of the Russian Academy of Science (Ioffe Institute). Papers and research adirections of their scientifc communities were examined. The results of numerical experiments confirmed the efficiency of the clustering algorithm used.

6 CONCLUSIONS

This paper describes the realization of monitoring the development of university scientific schools, which is one of functional components of the technological approach to university scientific knowledge management. Some models, methods, and technologies of university scientific knowledge life cycle support processes are considered.

The paper describes the developed model of a specialist which reflects the level of scientific activity productivity based on the calculation of entropy and overall scientific activity evaluation.

The approach to identification of university scientific schools based on the clustering of university scientific community by common interests is proposed.

The next stage of this work is to address the problem of assessment of university scientific activities and the degree of its integration with educational process.

The work was performed under grant "The development of an e-university's ontological knowledge base", state registration number 0213RK00305

REFERENCES

Klimov, S. M., 2002. Intellectual resources of society. SPb: IFEREL, Znanie.

Ackoff, R.L., 1989. From Data to Wisdom. In Journal of Applied Systems Analysis, vol. 16. pp. 3-9.

Zaim H., 2007. Performance of Knowledge Management Practices: a causal analysis. Knowledge Management, vol. 11(6), pp. 54-67. url:http://dx.doi.org/10.1108/13673270710832163.

Miles, I., 2005. Knowledge intensive business services: Prospects and policies. Emerald Group Publishing Limited, vol. 7(6). pp. 39-63. DOI 10.1108/14636680510630939.

Scarso, E., Bolisani, E., 2010. Knowledge-based strategies for knowledge intensive business services: A multiple case-study of computer service companies . Electronic Journal of Knowledge Management. url: www.ejkm com, vol. 8(1). pp. 151-160.

Tuzovskiy, A. F., 2007. The development of knowledge management systems based on a single ontological knowledge base. In Bulletin of the Tomsk Polytechnic University, vol2(310), pp. 182-185.

Sveiby, Karl-Erik., 1989. The Invisible Balance Sheet. Stockholm, 138 p.

Staab S., Schunurr H-P., Studer R., Sure Y., 2001. Knowledge processes and ontologies. IEEE Intelligent Systems, Special Issue on Knowledge Management, vol. 16(1), pp. 26-34.

NC STI RK. Scientific schools and priorities for the development of the country. URL: http://exclusive.kz/bez-rubriki/22068.

Trubina, I. O., Zabelina, I.N., 2011. Creating personnel's positive motivation in higher education institutions in the process of education and development of scientific schools. In Creative Economy, vol. 1(49), pp. 30-36.

Tuzovskiy, A. F., 2007. Creating and using a knowledge base of specialists' competence profiles at organizations. In Bulletin of the Tomsk Polytechnic University, vol. 310(2), pp. 186-189.

Zhomartkyzy, G., Balova, T., Milosz, M., 2014. Information Models and Methods of the University's Scientific Knowledge Life Cycle Support. Proceedings of the 2014 Federated Conference on Computer Science and Information Systems, pp. 279 – 285.

All-Russian Institute of Scientific and Technical Information, url: http://scs.viniti.ru/rubtree/main.aspx?tree=RV.

Adamic, L., Zhang, J., Bakshy, E., Ackerman, M.S., 2008. Knowledge sharing and yahoo answers: everyone knows something. Proceedings of the 17th international conference on World Wide Web. pp. 21-25.

Baesso, P. T.,Wolfgand, Matsui Siqueira S. and Cristina Vasconcelos de Andrade L., 2014. Finding Reliable People in Online Communities of Questions and Answers.-AnalysisofMetricsandScopeReduction. In Proceedings of the 16th International Conference on Enterprise Information Systems, pp. 526-535, DOI: 10.5220/0004954005260535.

Cantador, I., Castells P., 2011. Extracting multilayered Communities of Interest from semantic user profiles: Application to group modeling and hybrid recommendations . Computers in Human Behavior, vol. 27(4), pp. 1321-1336.

Bolshakova, E. I., Klyshinsky, E. S., Lande, D. V., Noskov, A. A,. Peskova, O. V., Yagunova, E. V., 2011. Automatic processing of natural language texts and computational linguistics: Textbooks. - M.: MIEM, 272 p.

Marmanis, H., Babenko, D., 2011. Algorithms of Intellectual Internet. Best practices for collecting, analyzing and processing data. Trans. from English. - SPb .: Symbol-Plus, 480 p.

Ioffe Physical-Technical Institute url: http://www.ioffe.ru/