Intermediate Models Evaluation for Ontology Development for Smart Spaces Technologies

Natalya Gusarova, Galina Artemova, Igor Kotciuba
Saint Petersburg National Research University of Information Technologies, Mechanics and Optics
Saint Petersburg, Russia
natfed@list.ru, {g.o.artemova, igor.kotciuba}@gmail.com

***Abstract*—This paper considers ontology evaluation as an important stage of the ontological engineering for smart spaces technologies. It gives an overview of the ontology evaluation methods and quality metrics for their classification. It also discusses the approach to ontology development using intermediate models (mind maps and concept maps) and key metrics for these models. We present a smart objects' ontology evaluation model and an algorithm for automated ontology development based such intermediate models as mind maps and concept maps using key metrics for their evaluation.**

I. Introduction

Intelligent systems connected with smart space technologies often use ontologies as a perspective method of structuring, development, integration smart objects. Study of ontology quality is an important aspect of ontological engineering. Automated ontology development requires formal methods of evaluation and optimization. Relevance of research in this area is confirmed by practical necessity, since the use of well-designed ontologies improves the efficiency of Semantic Web services, as well as by the fact that currently there are a large number of different approaches to assess the developed ontologies. As a result, there are problems of choosing an adequate technique of a series of developed ones, the application of which would be useful in solving a certain range of tasks, as well as developing new methodologies for assessing additional aspects of ontology quality, so research in this area continues to be valid.

II. Related works

A number of studies have reviewed the existing approaches to assessing the ontology quality. Since there are many options to choose bases for classification, different classification methodologies for assessing the ontology quality can be offered.

Proposed on the basis of ISO/IEC 25000:2005 (a standard for Software product Quality Requirements and Evaluation known as SQuaRE ISO (2005)) [1] automated framework for supporting the ontology assessment processes [2] considers such quality characteristics of ontologies as:

1) structural, reliability, functional adequacy, performance efficiency;
2) transferability, compatibility, maintainability;
3) operability, quality in use.

In addition, as shown in this study, before supporting the ontology assessment processes, OQuaRE should be improved in various ways. Namely, framework expansion with the quality requirements module is required, which would enable to determine the potential context of use; it would be useful for ontology evaluators. It is noted that each subcharacteristics of ontology quality has a set of metrics for quality assessment, and new methods can be determined by adaptation of this set of metrics for subcharacteristics associated with a set of quality characteristics, according to a SQuaRE-based quality model.

The following classifications of metrics for ontology quality evaluation are distinguished among others:

1) work [3] presents a set of metrics to evaluate ontology quality; the following metrics are distinguished: syntactic, semantic, pragmatic, and social;
2) works [4], [5] present the problem of ontology quality assessment based on modularization approach, indicating a set of characteristics that should be incidental to the ontology. These characteristics include authorization, encapsulation, self-containment, reusability, scalability, loose coupling, reasoning support. It is noted that during assessment of modular ontology quality, cohesion and coupling become important measures; modular ontology formalisms are also considered as means to support a set of characteristics. Work [5] postulates that ontology development "ab initio" requires considerable efforts and time, and the problem can be supported by the use of modularization approach;
3) work [6] presents metrics for evaluating both the schema, and the entire knowledge base of an ontology, including such metrics as class richness, attribute richness, inheritance richness, and relationship richness;
4) work [7] presents measures to assess such ontology aspects as functional, structural, usability measures.

Article [8] distinguishes both a number of objectives of the existing approaches to assessing the ontology quality, such as:

1) completeness and accuracy of dictionary of the considered domain;
2) structure adequacy in terms of taxonomy, relations;
3) performance in applications;
4) selection of the best ontology from the available set;

5) perceptibility from a cognitive point of view and options for ontology evaluation at different stages of development and use of ontologies, classification of these methods in the degree of automation, objects for analysis, and tools to determine the quality and maturity of ontologies.

According to the automation degree of various methods of ontology evaluation, manual, semi-automated, and automated groups of ontology assessment methods can be distinguished. In this regard, attention is paid to different aspects of supporting the activity of ontology evaluators. The use of different automatically computed metrics will significantly reduce expert's costs for their evaluation. Thus, according to the authors of [2], the use of new techniques by adapting the metrics for ontology quality subcharacteristics better coincides with the experts' wishes, since different combinations in different contexts of ontology application can be obtained. The use of cognitive ergonomics [8] will enable the expert to evaluate ontology at a glance, reducing the time for its assessment, and these techniques can be used in conjunction with another set of methods.

Issues of specific tasks of a particular domain and tasks of ontology development "ab initio" lead to the solution of ontology modularization questions [5]. Solving of these problems is inextricably linked to the issues of quality assessment in ontological engineering, since professionals-engineers need to compare different modularization methods and select the appropriate one for ontology modularization procedure. The paper also notes that it is necessary to pay attention to various aspects of modularization methods assessment, such as tool performance, data performance, and usability. Special attention is also paid to the fact that the choice of such a method must meet the specific requirements of an individual domain (environment).

The process of ontology development "ab initio" can also be supported by the use of intermediate models, the detailed elaboration of which will enable design of better ontologies. Let us consider the most commonly used ones.

During ontology development, much attention is paid to the visualization problems. Work [9] presents the basic visual models and methods in the development of ontologies and similar structures. One of the promising methods of this type are mind maps and concept maps that are widely used nowadays for visualization of ontologies at the design stage [10], when discussing the ontology structure [11].

Mind maps proposed by psychologist T.Buzan are widely used in various fields of human activity as a means of visualization, structuring, classification of ideas, to assist in training, problem-solving, decision-making [11].

Model of concept maps, giving more opportunities for formalization, can become another intermediate model in ontology development. It is noted in [12] that concept maps can be used as the first step in ontology constructing methods, being a means of expression for the expert and helping him to refine the structure of knowledge. It is emphasized that this method helps to reduce certain difficulties in the development of large-scale ontologies.

One of such environments, with a number of specific requirements, is the sphere of smart spaces technologies and ontologies designed for these purposes [13-17]. Work [13] pays special attention on using of ontologies in the field of searching, discovering, sharing information about residential environment services. Application development framework based on a domain ontology is used for discovering smart objects and their integrating in complex scenarios. Authors emphasize that the using of ontologies helps to understand the relations between sensors and actuators in smart entities and helps to access and compose them.

Work [14] highlights that ontologies are used for smart space application development for the input and output commands' data. However different data can have different input formats so it is noted that new execution components based on ontologies and ontology versions should be developed.

Another work [15] postulates that the use of ontologies in the area of intelligent transportation systems can allow generating smart clients and reusing of domain knowledge. It is shown that a systematic methodology should be proposed to develop ontologies with different algorithms for representation of data structure and data extraction.

Thus, the survey has showed that the process of ontological engineering is much facilitated using a variety of metrics for ontology evaluation, and the choice of an appropriate methodology for ontology assessment or the development of new approaches to their evaluation greatly depends on the assessment goals and the context of ontology use. One of the environments possessing the specific character is the sphere of smart spaces, to support the different tasks of which ontologies are applied.

This paper considers a model of ontology evaluation in the field of smart spaces, which in accordance with the existing classifications can be interpreted as:

1) goal: perceptibility, performance in applications;

2) analysis object: structure of intermediate models of ontology development – mind maps and concept maps;

3) analysis tool: topology analysis of ontology graph;

4) degree of automation: semi-automatic (after the automatically computed metrics values, at the stage of intermediate model optimization the expert completes the analysis by himself);

5) application stage: development and prototyping, testing before release and implementation of ontology.

III. METHOD

A. Metrics' selection

Algorithm for automated ontology development requires consideration of intermediate models' specifics. Works [9, 18] highlight some algorithms for constructing mind maps and

concept maps. Work [18] notes the connection between the construction of mind maps and concept maps when developing ontologies. It is postulated as well that this algorithm can be used in various tasks of ontology constructing, when environment expert works with abstract concepts. However, the paper does not pay attention to an incremental algorithm for mind maps constructing. Work [9] draws attention to the fact that to construct a mind map one can use a modified five-step algorithm for visual construction of ontologies as a model conceptually describing the environment, but it does not consider the algorithms of proceeding from mind map construction to a concept map in further ontology development. Moreover, these papers focus on the immediate expert's work on optimization of intermediate models, and methods for their automated optimization are not considered.

Particular attention in this algorithm should be paid to optimization at the intermediate stages, namely, at the stage of mind map optimization and concept map optimization, to avoid additional errors in ontology development. In addition, automated optimization of intermediate models related to visual means of knowledge processing will improve their quality and reduce expert's costs when working with them.

Metrics selection for mind maps and concept maps optimization depends on their structural specifics. Informative analysis of characteristics of mind maps as graph structures defined their main characteristics:

1) have a tree structure;
2) tend to have a small number of concepts of the first level (not exceeding 4-5);
3) concepts of the same hierarchy level have the same type of relationship with the parent concept;
4) have a property of uniformity – the difference in the number of levels of various branches should not exceed 2;
5) the number of child concepts should not exceed 7±2;
6) branch depth should not exceed 7±2.

Main characteristics of concept maps as graph structures are as follows:

1) include relationships between the concepts;
2) at the development stage require in-depth analysis of structural interactions between certain concepts in the domain.

It should be also noticed as a characteristic of concept maps that they can represent a tree, but not necessarily.

Comparative evaluation of the effectiveness of existing metrics to optimize mind maps and concept maps was carried out using multi-criteria analysis methods [20, 21], which enabled to identify the following set of metrics to optimize mind maps:

1) metrics of depth, diameter, height, concept height (layer), graph width (including absolute, average, maximum);
2) metrics of tree analysis: metrics of perfect balance and AVL tree balance (properly organized tree is a perfectly balanced tree in which for each tree node the number of nodes in the left subtree differs by at most 1 from the number of nodes in the right subtree. Tree is AVL-balanced (or simply "balanced") if for each its node height of the left subtree differs at most by 1 from height of the right subtree);
3) out-degree metrics for analyzing the number of edges emanating from any concept of mind map, starting from the central image;
4) cycle metrics (because tree is a graph without cycles);
5) metrics of graph branching measurement: enables to evaluate the "distribution" of graph nodes, which have leaves and non-leaf nodes among children.

These metrics enable to evaluate the quality of the constructed mind map, and to support the process of knowledge representation on a predetermined topic, avoiding possible moving away from it: for example, the balance metrics ensures that the expert will detail the central concept of the domain, not being carried away with specification of other, particular concept.

To optimize concept maps, analysis revealed the following metrics:

1) metrics of depth, diameter, height, concept height (layer), graph width (including absolute, average, maximum);
2) metrics of tree analysis: metrics of perfect balance and AVL tree balance. These metrics can be used when the concept map is a tree; however, this is not a necessary condition for this intermediate model;
3) metrics of in- and out-degree: unlike mind maps, where only one edge comes in each concept, except for the central one, in concept maps this parameter, as out-degree, may have different values;
4) cycle metrics (presence of cycles prevents perception; this metrics also enables to evaluate whether the conceptual map can be decomposed into a tree or not);
5) metrics of graph branching measurement (as in the case of mind maps);
6) edge density [22] characterizing graph closeness to a complete graph (clique). This metrics allows automatic detection of errors in the expert's concept map, since clique is not a typical situation when denoting concept-"parents" and "children";
7) metrics of diverse number of links – since the presence of different types of relationships between concepts is one of the properties of a concept map by definition;
8) metrics of graph complexity, including analysis of the ratio of nodes with multiple inheritance in respect of the number of all graph nodes and analysis of the average number of parent nodes at the graph node is

also applicable for concept map evaluation due to the presence of these characteristics;

9) solving the problem of in-depth analysis of structural interactions between the concepts can be supported by the methods of connectivity analysis of graph structures. Unlike the traditional studies of graph connectivity, q-analysis [23, 24] method allows for a more thorough judging on the system connectivity, establishing the presence of cross-impact of system simplexes through links between them. Q-analysis of system connectivity reveals simplexes having maximum influence on the processes in system, as well as nodes, which are more rational to be chosen as controlling; it becomes possible to trace the effect of various local changes on other system elements and on the system structure in general.

Table I presents a summary of metrics of intermediate models when developing ontologies (mind maps and concept maps):

TABLE I. LIST OF METRICS FOR AUTOMATED OPTIMIZATION OF INTERMEDIATE MODELS WHEN DEVELOPING ONTOLOGIES

Intermediate model / **Metric**	**Mind map**	**Concept map (tree/ not tree)**
1. Graph depth (absolute, average, maximum)	+	+
2. Graph diameter	+	+
3. Graph height	+	+
4. Concept height (layer)	+	+
5. Graph width (absolute, average, maximum)	+	+
6. Tree balance	+	+/–
7. Perfect tree balance	+	+/–
8. In-degree	–	–/+
9. Out-degree	+	+
10. Cycle metrics	+	+
11. Metrics of graph branching measurement	+	+
12. Edge density	–	+
13. Metrics of number of links diversity	–	+
14. Metrics of graph complexity	–	+
15. Simplicial q-analysis of graph connectivity	–	+

B. Evaluation's algorithm

Algorithm for ontology development support based on mind maps and concept maps using the proposed metrics is presented below:

1) the central concept is determined at the first step; future work will be linked to its specification;

2) formation of mind map, which represents structured knowledge:

- environment glossary identification;
- formation of links between its concepts ("bottom-up" grouping procedure, determination of categories, metaconcepts);
- visual representation of hierarchical levels in the system of selected concepts (specification by "top-bottom" principle, from the central concept to the concepts on certain lower levels);
- further specification and clarification of metaconcepts if necessary;
- mind map optimization:

1) addition, resolving the problems of synonymy, contradictions, redundancy;

2) mind map assessment using metrics (see metrics in Table I);

3) visual representation of a balanced mind map;

4) mind map coordination with an expert;

5) concept map formation with the division of class and individual concepts, types of relations between them:

- determination of connections (relationships and interaction) between the concepts;
- optimization of concept map together with an expert (replacement of synonymous relations, removing redundancy):
- addition, resolving the problems of synonymy, contradictions, redundancy;
- concept map assessment using metrics (see metrics in Table I);

6) division into meaningful sections, final critical replenishments, and corrections;

7) results recording in a document in ontology language (e.g., OWL format).

IV. RESULTS AND DISCUSSION

Algorithm for ontology development support based on mind maps and concept maps using the proposed metrics is presented below using activity diagram (Fig. 1).

Examples of computations (metrics' numbers from Table I) were performed on real concept maps [25], [26] developed for the training process that can be used in the field of smart space for learning as a perspective trend of learning management systems' development using ontologies [27].

Interaction of program objects regulated on time is presented below using sequence diagram (Fig.2).

Table IV presents metrics from Table I using scale of priority that allows compare concept maps from Fig.3 and Fig.4. Designation «>» means bigger priority, designation «<»- smaller priority, designation «=» - equal priority.

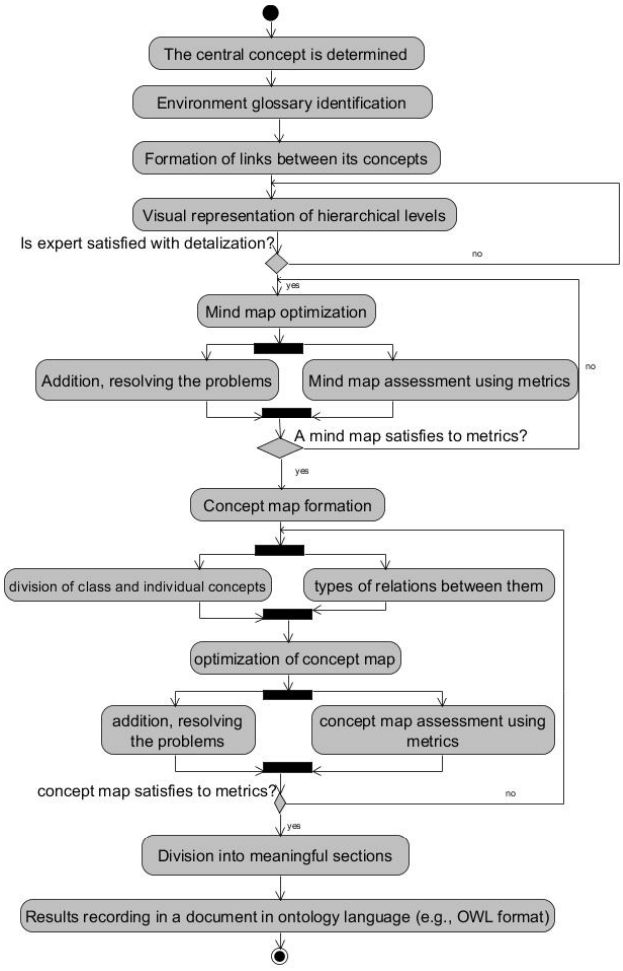

Fig. 1. Algorithm for ontology development support based on mind maps and concept maps using metrics

Received values can give such recommendations to an expert that takes part in ontology development in smart spaces technologies as:

- delete map's concepts if their amount is excessive that can facilitate perception;
- delete or add map's concepts if it is like a tree that can improve balance of a concept map;
- delete or add map's concepts if it is not a tree that influence simplexes' connectivity in a concepts' complex.

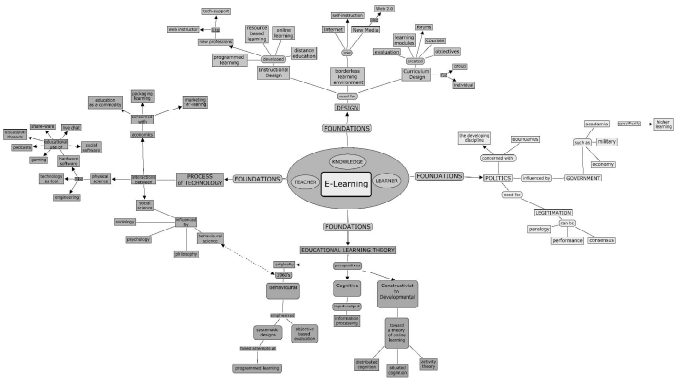

Fig. 3. Concept map "E-learning" (tree)

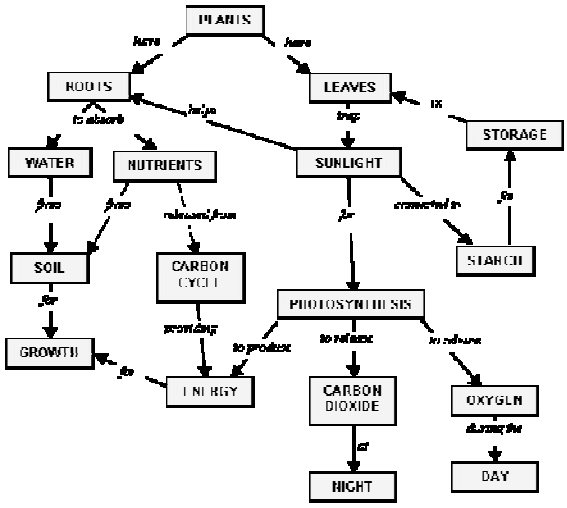

Fig. 4. Concept map "Photosynthesis" (not tree)

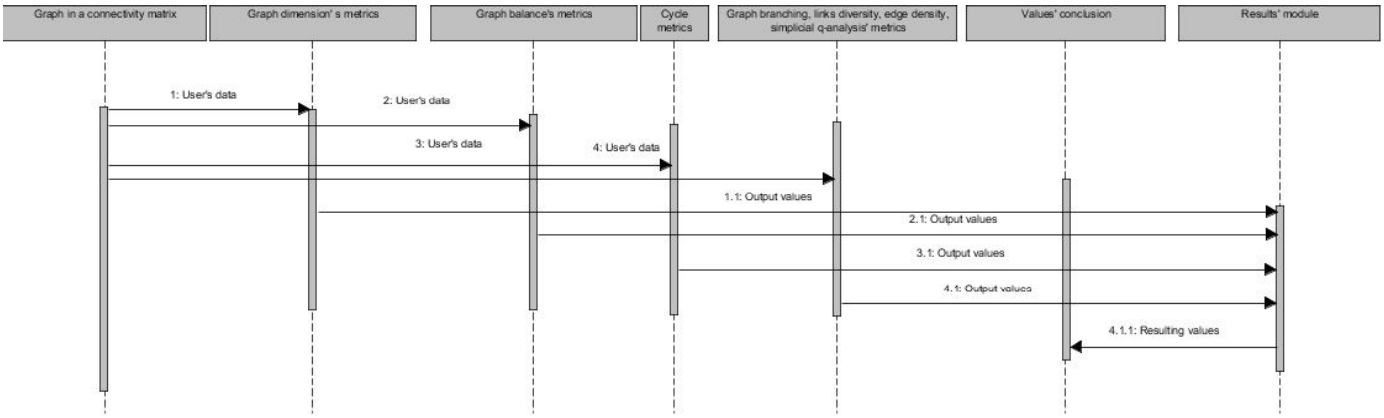

Fig. 2. Sequence diagram for ontology development support based on mind maps and concept maps using metrics

TABLE II. VALUES FOR CONCEPT MAP (TREE)

Metric's number	1	2	3	4	5	6	7
Value	146 3.24 5	9	6	1-6	67 11.18 36	10	No
8	9	10	11	12	13	14	15
-	0-6	0	0,35	0,03	0,31	0	ecc: ∞

TABLE III. VALUES FOR CONCEPT MAP (NOT TREE)

Metric's number	1	2	3	4	5	6	7
Value	14 4.67 5	5	6	1-6	17 2.83 5	-	-
8	9	10	11	12	13	14	15
0-2	0-3	0,24	1	0,17	0,83	1,24	ecc: 0-2, ∞

TABLE IV. SCALE OF PRIORITY

Metric	Concept map "E-learning"	Concept map "Photosynthesis"
Graph depth (average)	>	<
Cycle metrics	>	<
Metrics of number of links diversity	>	<
Metrics of graph complexity	>	<
Graph depth (absolute)	<	>
Graph diameter	<	>
Graph width (absolute, average, maximum)	<	>
Metrics of graph branching measurement	<	>
Graph depth (maximum)	=	
Graph height	=	
Concept height (layer)	=	
Out-degree	=	
Edge density	=	

Special software developed for analysis of mind maps and concept maps allows receive values of all metrics from table 1. The example of one of metrics (simplicial q-analysis of graph connectivity) is given below:

```
int q1, q2, simplicial; //eccentricity value of each concept
    float ecc;
    for (i = 0; i < n; i ++) {
      q1 = numbl[i] - 1;
      q2 = -5;
      for (j = 0; j < n; j ++) {
        if (i != j) {
          simplicial = numb_same(matrix [i], matrix [j],
numb [i], numb [j]) - 1;
          if (q2 < simplicial) {
            q2 simplicial;
          }
        }
      }
      cout << i + 1 << " - ";
      if (q2 == -1) {
        cout << "infinity; ";
      } else {
        ecc = (q1 - q2) / (q2 + 1);
        cout << ecc << "; ";
      }
    }
```

V. CONCLUSION

Thus, this paper substantiates intermediate models evaluation for ontology development in smart spaces technologies, proposes a smart objects' ontology evaluation model as well as an algorithm for automated ontology development based on intermediate models using the proposed metrics. The considered model was realized as special software and examples of computations were performed on real concept maps.

REFERENCES

[1] ISO/IEC 25000:2005.

[2] Duque-Ramos, A., Fernandez-Breis, J.T., Iniesta, M., Dumontier, M., Aranguren, M.E., Schulz, S., Aussenac-Gilles, N., Stevens, R. 2013. Evaluation of the OQuaRE framework for ontology quality. *Expert Systems with Applications*, 40 (2013), 2696-2703.

[3] Burton-Jones, A., Storey, V., Sugumaran, V., Ahluwalia, P. A semiotic metrics suite for assessing the quality of ontologies, *Data & Knowledge Engineering*, v.55 n.1, p.84-102, 2005.

[4] Wang, Y., Bao, J., Haase, P. and Qi, G., Evaluating formalisms for modular ontologies in distributed information systems. 2007. *Lecture Notes in Computer Science*, 2007.

[5] Sunjoo, Oh, Heon, Y. Yeom. 2012. A comprehensive framework for the evaluation of ontology modularization. *Expert Systems with Application*, 39 (2012), 8547-8556.

[6] Tartir, S., Arpinar, I., Moore, M., Sheth, A., Aleman-Meza, B., OntoQA: metric-based ontology quality analysis, in: Proceedings of the Workshop on Knowledge Acquisition from Distributed, Autonomous, Semantically Heterogeneous Data and Knowledge Sources (KADASH), Citeseer, 2006.

[7] Gangemi, A., Catenacci, C., Ciaramita, M., Lehmann, J., Modelling ontology evaluation and validation, Proceedings of the 3rd European conference on The Semantic Web: research and applications, June 11-14, 2006, Budva, Montenegro.

[8] Gavrilova, T., Gorovoy, V., Bolotnikova, E., Gorelov, V. Subjective metrics for ontology evaluation. In Proceedings of the International Conference "Knowledge-Ontologies-Theories"(ZONT-2009). Institute of Maths CO RAN, 2009, 178-186.

[9] Gavrilova, T., Leshcheva, I., Strahovich, E. 2011. About use of visual concept models in teaching. *Vestnik of Saint-Petersburg State University*, 4 (2011), 124-150.

[10] Gavrilova, T., Gladkova, M. 2014. Big data structuring: the role of visual models and ontologies. *Procedia Computer Science*, 31 (2014), 336-343.

[11] Roussey, C., Pinet, F., Kang, M.A., Corcho, O. Ontologies in Urban Development Projects, Advanced Information 9 and Knowledge Processing 1, Web: 10.1007/978-0-85729-724-2_2.

[12] Starr, R.R., Oliveira, J.M.P. 2013. Concept maps as the first step in an ontology construction method. *Information Systems*, 5 (2013), 771-783.

[13] Albreshne, A., Lahcen, A.A., Pasquier, J. 2014. Using a Residential Environment Domain Ontology for Discovering and Integrating Smart Objects in Complex Scenarios. *Procedia Computer Science*, 32 (2014), 997-1002.

[14] Palviainen, M., Kuusijärvi, J., Ovaska, E. 2014. A semi-automatic end-user programming approach for smart space application development. *Pervasive and Mobile Computing*, 12 (2014), 17-36.

[15] Gregor, D., Toral, S., Ariza, T., Barrero, F., Gregor, R., Rodas, J., Arzamendia, M. 2015. A methodology for structured ontology construction applied to intelligent transportation systems. *Computer Standards & Interfaces* (2015).

[16] Balandin, S., Waris, H. Key Properties in the Development of Smart Spaces in Volume 5615 of the series Lecture Notes in Computer Science, *Human Computer Interaction International conference* (HCII 2009). San Diego, USA, July 2009. pp. 3-12.

[17] Smirnov, A., Kashevnik, A., Shilov N., Balandin S., Oliver, I., Boldyrev, S. On-the-Fly Ontology Matching in Smart Spaces: A Multi-Model Approach in Volume 6294 of the series Lecture Notes in Computer Science, *3rd Russian Conference on Smart Spaces – ruSMART 2010*. St-Petersburg, Russia, August 2010. pp. 72-83.

[18] Mouromtsev, D.I. Development of a complex ontology of optics. CEUR Workshop Proceedings. 2012. Vol. 849. pp. 12p.

[19] Gavrilova, T., Leshcheva, I., Kudryavtsev, D. Use of knowledge engineering' models for training of specialists in sphere of IT. 2012. *System programming*, Saint-Petersburg State University, 1 (2012), 90-105.

[20] Pawlak, Z. 2002. Rough sets and intelligent data analysis. *Information Sciences*, 147 (2002), 1-12.

[21] Saaty, T. L. Decision Making for Leaders. *RWS Publications*, 4922 Ellsworth Ave, Pittsburgh PA, 1999.

[22] Karpenko, A. 2010. Measures of importance of concepts in semantic net of ontological knowledge base. *Science and education*, 7 (2010), 1-12.

[23] Casty, J. Connectivity, Complexity and Catastrophe in Large. *Scale Systems*. N.Y., L.: Willey-Interscience, 1979.

[24] Gorelova, G., Melnik, E. 2008. About opportunity and synthesis of structures of fault-tolerant distributed management information systems based on cognitive approach. *Artificial intelligence*, 2008, 638-648.

[25] ETEC 511: mapping foundations of ed-tech and e-learning. Web: https://erinbgillespie.wordpress.com/theme/plants/etec-511-mapping-foundations-of-e-learning/

[26] Kinchin, I.M. 1998. Constructivism in the classroom: mapping your way through. In Proceedings of the British educational research association annual research student conference The Queen's University of Belfast, August 26th - 27th 1998. Web: http://www.leeds.ac.uk/educol/documents/000000811.htm

[27] Simon, B., Miklos, Z., Nejdl, W., Sintek, M., Salvachua, J. Smart Space for Learning: A Mediation Infrastructure for Learning Services. *In Proceedings of the Twelfth International Conference on World Wide Web*, Budapest, Hungary, 20-24 May 2003.