The Methodology for Ontology Development in Lesson Plan Domain

Aslina Saad

Computing Department
Faculty of Art, Computing & Creative Industry
UPSI
Tanjung Malim, Malaysia

Shahnita Shaharin

Computing Department
Faculty of Art, Computing & Creative Industry
UPSI
Tanjung Malim, Malaysia

Abstract—**Ontology has been recognized as a knowledge representation mechanism that supports a semantic web application. The semantic web application that supports lesson plan construction is crucial for teachers to deal with the massive information sources from various domains on the web. Thus, knowledge in lesson plan domain needs to be represented accordingly so that the search on the web will retrieve relevant materials only. Essentially, such retrieval needs an appropriate representation of the domain problem. The emergence of semantic web technology provides a promising solution to improve the representation, sharing, and re-use of information to support decision making. Thus, the knowledge of lesson plan domain needs to be represented ontologically to support efficient retrieval of semantic web application in the domain of lesson plan. This paper presents a new methodology for ontology development representation of lesson plan domain to support semantic web application. The methodology is focused on the important model, tools, and techniques in each phase of the development. The methodology consists of four phases, namely requirements analysis, development, implementation, evaluation and maintenance.**

Keywords—knowledge representation; methodology; ontology development; lesson plan

I. INTRODUCTION

Ontology is widely used for knowledge representation in artificial intelligence, information retrieval and semantic web [15]. Ontology provides a common understanding of specific domains and is also expressed as a formal representation of knowledge by a set of concepts within a domain and the relationship between these concepts [18]. [5] has built an ontology of lesson plan in the form of hierarchical taxonomy that shows the semantic relationships between terms in lesson plan domain. The taxonomy was then used to produce relevant search results using semantics approach in a case based reasoning (CBR) system. This is parallel to what was stated by [8], ontology can enhance CBR systems in many dimensions. Related terms are looked up from the lesson plans ontology, a structured data source. An evaluation study was designed to examine the effectiveness of the system using this representation and have shown positive results. However, the search result is limited to cases defined in the database.

Past research shows that lesson planning imposes significant burden on teachers and causes excessive workload among teachers as they need to spend a lot of time to prepare their lessons. Several efforts have been carried out to overcome this predicament, including the development of a web-based system to assist teachers in such a task.

Various platforms have been established to enable the sharing of information among teachers, such as web pages, blogs, online systems, Slideshare[1], and even social networking sites such as Facebook[2]. However, searching relevant materials or contents using such multiple platforms will result in too much information being fed to teachers. Consequently, teachers need to filter that information to select materials that really meet their teaching needs. Invariably, such effort is laborious and taxing, which further burdens their teaching workload.

Furthermore, almost all of the web-based applications use attribute-value representation and they are stored in databases, which need constant updating and verification by the system administrator. Arguably, the uses of such databases have both benefits and limitations. According to [13], the usefulness of the database management system based on the three models, namely hierarchical, network and relational models, is severely restricted by the failure to take into account the semantic of databases.

The above limitation can be overcome by applying semantic web technology. In particular, such semantic system can be used for inter organization data sharing and reuse. The purpose of this system is to let information on the Internet to have richer semantics in order to facilitate computers to determine information that is important and relevant to various users' needs, thus improving the interoperation among the entities on the Internet.

In essence, the semantic web technology is capable to connect a particular website to other websites through the use of knowledge representation. Each site in the network of internet is connected to each other by an existing relationship that has been defined in terms of knowledge representation. Thus, the search for online information can be implemented more intelligently by focusing on the relevant domain. Given this capability, knowledge representation can be applied to a dedicated semantic web application to help diverse users, especially teachers in lesson planning.

Knowledge representation is important to produce intelligent systems based on knowledge as a key element to enable the process of reasoning and decision making. According to [2], knowledge representation is one of the core

---

[1] http://www.slideshare.net/
[2] https://www.facebook.com/

elements in the field of artificial intelligence, which is an important aspect of problem solving. Such focus on problem solving based on knowledge representation is also stressed by [16], who state that a computer system that is capable of performing tasks that require human intelligence entails such representation.

Essentially, the purpose of understanding what knowledge is and what are the types of knowledge that exist allows us to use it in artificial systems [6]. Thus, in the context of lesson plan construction, the issue of deciding what is to be stored and how memory should be organized in order to retrieve and reuse previously prepared teaching plans effectively and efficiently needs to be addressed urgently.

In general, many different architectures have been used for knowledge representation, including ontology. According to [3], ontology is widely considered as a promising approach for capturing and representing knowledge. On one hand, [20] asserts that ontology as an explicit formal specification of a shared conceptualisation. On the other hand, according to [16], ontology is a method that defines terms which are commonly accepted for a particular domain to enable the effective sharing of information among researchers. In essence, this definition encompasses the concepts and their relationships in that domain.

Accordingly, ontology will help ensure that the terms and symbols used are defined with clear intention. Moreover, computing is the key component that enables logic and ontology-based representation to be implemented in a computer program. In this regard, [22] describe that the semantic web ontology language should include five key components, namely the concept, taxonomy, relations and functions, axioms, and instances. Particularly, the concept involves the explanation on common issues, attributes, and facets.

II.   EXISTING METHODOLOGIES

The literature is replete with studies in which several scholars and researchers have proposed several ontology development methodologies. According to [19], here are two ways of conceiving ontology construction, the bottom up and top down approach. Such scholars include [3], [14], [12], and [10], where the third and the forth researchers introduced ontology development models called *Methontology* and *Model-based and Incremental Knowledge Engineering* (MIKE), respectively. In essence, each methodology comprises several phases, which are classified and contrasted as summarized in Table I.

Clearly, there are some methodologies that are more comprehensive compared to others, such as [12] model, that also focus on phases of evaluation, documentation and maintenance. In contrast, [3] model does not emphasize on the evaluation phase prior to the maintenance phase. In addition, the development phases of [14] model and [10] terminate at the implementation phase, without the evaluation, documentation, and maintenance phases.

TABLE I.   THE COMPARISON AMONG EXISTING ONTOLOGY DEVELOPMENT MODELS

Ohgren and Sandkuhl (2005)	Uschold and King (1995)	Methontology (Fernandez, 1997)	MIKE (Angele, 1998)
Requirement analysis	Specification	Specification	
		Knowledge acquisition	Elicitation
Development	Conceptualization	Conceptualization	Interpretation
	Formalization	Formalization	Formalization
		Integration	Design
Implementation	Implementation	Implementation	Implementation
Evaluation Maintenance		Evaluation	
		Documentation	
		Maintenance	

Interestingly, the model proposed by [12] is more comprehensive as evidenced by its nine development phases as opposed to other development models that comprises of four or five development phases only. Essentially, [3] model is similar with the development methodology of any application domains. In their model, the phases involve requirement analysis, development, implementation, evaluation, and maintenance, which are more distinct in the computing realm. Such development phases can help developers to familiarise with the terms or nomenclatures that they are not used to in the ontology field, such as conceptualization, formalization, integration, and interpretation. The development model introduced by [3] consists of requirement analysis, development, implementation, evaluation, and maintenance phases. Fig. 1 summarizes the four development phases, together with their end results.
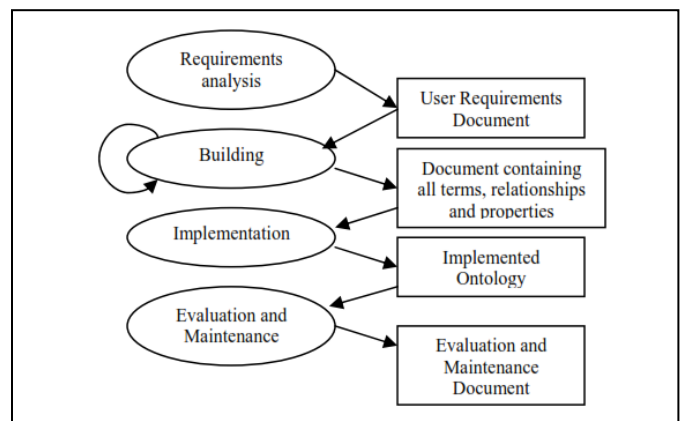

Fig. 1.   Ontology development phases from [3]

Apparently, all the above methodologies emphasize requirement analysis as the first phase although different terms have been used. This phase is then followed by ontology development based on the acquired knowledge. Later, the developed ontology is then implemented and finally evaluated.

III.   APPLIED METHODOLOGY FOR LESSON PLAN ONTOLOGY DEVELOPMENT

Based on the comparison of the above methodologies, a new methodology was formulated based on [3] and [12] models to serve as a guideline for knowledge representation methodology as summarized in Table II.

A. Phase I: Requirement analysis

Requirement analysis is the first phase of the methodology that comprises two sub phases, namely specification and knowledge acquisition.

TABLE II.   THE COMPARISON AMONG EXISTING ONTOLOGY DEVELOPMENT MODELS

Phase		Activities
Requirement analysis	Specification	Identifying ontology specification includes: • The purpose of the developed ontology • Target users of the ontology • Ontology usage scenarios • Scope of the ontology • User requirements • Requirements of equipment and software
	Knowledge acquisition	Acquiring informal information related to knowledge and problem-solving process of subject matter experts using surveys, structured interview, observation, document analysis, and structuring techniques.
Development	Conceptualization	Developing knowledge representation in a semi-formal format using graphical representation.
	Formalization	Changing the semi-formal knowledge representation to formal knowledge representation.
	Integration	Identifying any appropriate existing ontology that can be integrated into the ontology being developed.
Implementation	Implementation	Transforming human-readable representation into machine-readable representation.
Evaluation Maintenance	Evaluation and Maintenance	Evaluating and assessing the developed ontology in meeting the requirement specifications. Identifying individuals to update and maintain the developed ontology.

1) Specification: This phase involves identifying all specifications of ontology requirements, which includes the objectives, target users, usage scenarios, scope, needs, and requirements of equipment and software in the development process ontology. For example, such needs for equipment and software include yEd[3] Graph Editor (to construct the conceptual modelling), StarUML[4] (which is to generate UML) and Protégé 5.0[5] (as an ontology language) that will be used to develop ontologies in the implementation phase. Identifying the specifications of ontology include:

* *The purpose of the ontology*: The problem of the domain, involving the construction of lesson plans for semantic application.

* *The target users of the ontology*: Teachers of any levels, including trainee teachers, inexperienced teachers, or experienced teachers.

* *The ontology usage scenarios*: Information retrieval related to lesson plan construction. The scope of the ontology for daily lesson plan.

* *The user requirements*: To support information retrieval based on keywords that are inserted by users to the semantic application.

* *The requirements of equipment and software*: The software to support ontology development conceptually and physically. For example, yEd graph editor were used for modeling, (semantic net and class diagram) and Protégé 5.0 for formalizing the developed ontology.

2) Knowledge acquisition: The first phase in the acquisition of knowledge was the elicitation process. In this phase, the procurement process related to information of the selected domain was implemented. Various techniques can be done for this procurement including the use of concept maps as a means of expression for the expert [17]. Among the activities carried out in this phase was a review of related literature, an analysis of related documents, a survey, and a structured interview with domain experts. Information obtained from these activities was recorded in the form of a natural language representation that is human-readable. The following are the techniques of knowledge acquisition implemented in this study.

a) Survey: A preliminary study was carried out using a quantitative approach with the aim to get feedback from Malaysian teachers about their daily lesson planning. An online survey questionnaire was developed using Google Form[6] to facilitate faster distribution and administration involving a wider participation compared to the conventional survey questionnaire. Essentially, the procedure of the survey involved three main steps. First, the construction of questions was performed by focussing on the components of a lesson plan, factors that influence lesson plan preparation, and materials or resources for lesson planning. Second, a survey questionnaire was distributed online. Third, the collected survey data were analyzed using appropriate descriptive statistical method.

* *Respondents*: The sample size of the online survey was made up of 150 teachers consisting of 117 female teachers and 33 male teachers. These respondents came from a diverse background, who had teaching experiences ranging from one to 30 years, and they worked in several Malaysian schools across the nation.

* *Research Instrument*: The research instrument used in the survey consists of three parts, which are demographics, lesson plan preparation, and materials parts.

b) Interview: A structured interview method involving 10 respondents was carried out with the main aim of eliciting further information by asking several pertinent questions to

---

[3] https://www.yworks.com/products/yed
[4] http://staruml.io/
[5] http://protege.stanford.edu/

[6] https://docs.google.com/forms/

verify some of the findings from the survey. In addition, a teach back technique was embedded in the interview sessions, in which important components or elements of a lesson plan based on [4] finding was presented to the interviewees. In particular, they were required to rank each element according to its perceived importance to the retrieval of information. The questions comprise several aspects, namely the process of producing a lesson plan, references in preparing the lesson plan, factors that are considered vital in producing the lesson plan, problems faced by teachers in preparing the lesson plan, and the necessity for tools to help them perform the task with ease.

c) Document analysis: Lesson plan documents were collected from the teachers during the interview sessions other than online resources. The documents were compared and analysed, such as to identify standard elements in a a lesson plan and the tools used for lesson planning. The analysis revealed that some of the available tools to support lesson planning were SmartLP[7], INTIME[8], KITE[9], The Lesson Planner Lesson Planning System (LPS)[10], RPH Online[11], PlanBookEdu[12], Planboard[13], Common Curriculum[14], Core Lesson Planner[15], and PlanBook[16].

B. Phase II: Development

The construction phase was a repetitive process, occurring in a cycle. In essence, this phase comprised three sub phases, namely the conceptualization, formalization, and integration sub phases. In each cycle, an evaluation was performed, and any changes were implemented to improve the constructed representation. Further explanation of each sub phase is as follow:

1) Conceptualisation phase: In this sub phase, the conceptual model for a specific domain (i.e., the lesson plan) was developed using semantic net as shown in Fig. 2. The knowledge for the domain was represented using a graph consisting of vertices to denote objects, concepts, domain entity, and edges. And each edge is basically a line connecting two vertices. As the lesson plan ontology construction was carried out within Malaysian context, the terms shown in Fig. 2 to Fig. 6 were constructed in Malay language.

2) Formalisation phase: In this phase, the conceptual model was transformed into a semi-formal representation using Unified Modeling Language (UML), specifically a class diagram. This is shown in Fig. 3. Then, a formal representation was formed to construct the ontology of the lesson plan domain. The ontology editing package Protégé 5.0 was used to develop the ontology for Daily Lesson Plan domain as shown in Fig. 4. Fig. 5 illustrates examples of the instance of the constructed class.
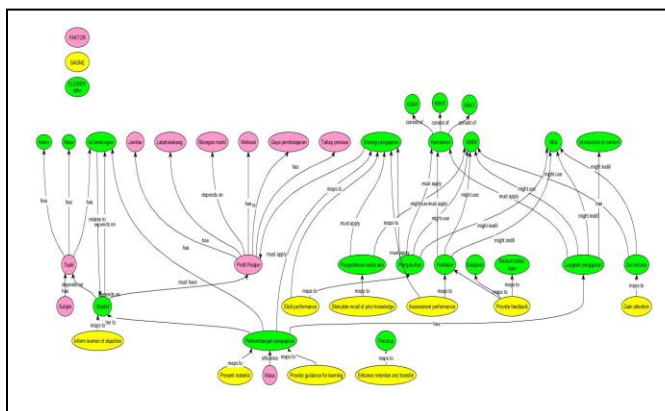

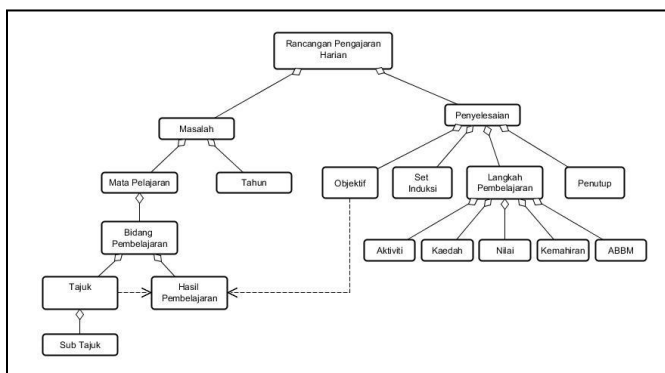Fig. 2. Semantic net for lesson plan domain

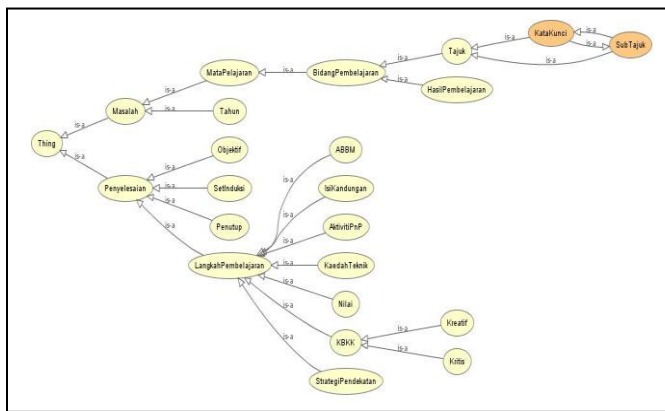Fig. 3. Class diagram for lesson plan domain

Fig. 4. Daily lesson plan ontology developed using Protégé 5.0

3) Integration phase: Any existing ontology for the lesson plan domain was identified in this phase by processing which parts of the ontology were appropriate or otherwise. If such ontology was suitable, it would be integrated into the developed ontology. A lesson plan ontology by [4] in a hierarchical form was compared to the constructed ontology. Then, these two ontologies of daily lesson plans domain were integrated during the development process.

---

[7] http://smartlp.upsi.edu.my/
[8] http://www.intime.uni.edu/casestudies
[9] http://kite.missouri.edu/
[10] https://www.oncoursesystems.com/products/detail/lessonplanner
[11] http://rphonline.teknologihijau.net/
[12] http://planbookedu.com/
[13] https://planboard.chalk.com/
[14] http://www.commoncurriculum.com/
[15] http://www.coreplanner.com/
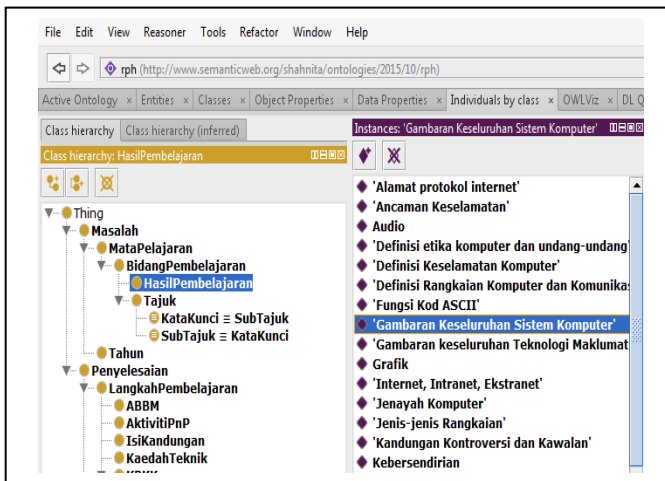[16] https://www.planbook.com/

Fig. 5.   Protégé 5.0 interface

C. Phase III: Implementation

The main aim of this phase is to change the human-readable representation to machine-readable representation. According to [7], RDF is a standard model for data interchange on the Web. Both RDFS and OWL are modeling languages for describing RDF data. For example, RDFS allows users to express the relationships among data by standardizing them using a flexible, triple-based format and then providing relevant vocabulary or keywords, such as "rdf:type" or "rdfs:subClassOf", which can be used to express such data. On the other hand, OWL is more powerful as it describes data models more efficiently using appropriate database queries and automatic "reasoners". Furthermore, OWL provides useful annotations to help transform the data models into the real world.

Such machine-readable representations include Ontology Web Language (OWL) format, which can be understood by the computer as illustrated in Fig. 6. In this study, Protégé 5.0 was used to convert the conceptual model to such representation.

Fig. 6.   The representation in OWL format

D. Phase IV: Evaluation and Maintenance

The third phase involved evaluating and assessing the developed ontology to determine whether it meets the requirement specifications which is to support retrieval. This is in line with was discussed by [22] which state that the organization of elements in knowledge representation must facilitate the retrieval of useful information. As a mean to prove the developed ontology in lesson plan domain, the constructed ontology was implemented in i-Rph[17] system, a semantic application for lesson planning.

This implementation involved the development of a system prototype using the developed knowledge representation. More importantly, the development of the system was based on the prototype development model comprising six (6) phases, namely preliminary study, requirement definition, system design, development and evaluation, implementation and maintenance. Fig. 7 shows all the six phases of the i-Rph system implementation.
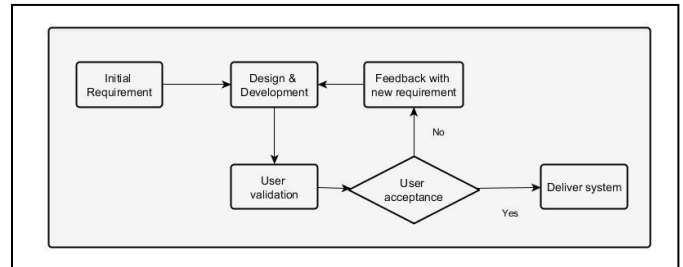

Fig. 7.   Evolutionary prototyping model (Adapted from [11])

This phase is challenging as, according to [9], information systems are not easy to be assessed, and there are many aspects to be considered in the assessment process. Moreover, with the emergence of new internet technologies, it is now more difficult to measure IS effectiveness. This is especially true given that internet provides a borderless, non-stop, and flexible communication medium.

Such evaluation will be performed using a quantitative approach by means of an experiment. This experimental approach is selected because it is one of the effective means in evaluating the implementation of a software system [1]. The evaluation phase will be carried out to evaluate the effectiveness of the constructed ontology based on the retrieval mechanism supported by the representation.

The evaluation of the system will be based on the Form Four's ICT curriculum as the scope of the study involves the Daily Lesson Plan (DLP) of the same subject matter. The respondents will be required to construct the lesson plan based on specific details, such as a particular topic or learning objectives. In constructing such lesson plan, they have to use three different applications: a) i-Rph which is the system developed using ontology representation, b) SmartLP system which uses attribute value representation, and c) any free, non-proprietary search engine. They are then required to answer a survey concerning aspects of information quality based on [21] IS Effectiveness Model.

The sample of this study will involve 20 trainee teachers from the Computing Department of the Faculty of Art, Computing, and Creative Industry, UPSI who will be undergoing teaching practicum at several secondary schools. These trainees will teach the Information and Communication Technology (ICT) subject at such schools. The selection of the sample will involve students majoring in computing because of their extensive exposure to the use of technology in education.

The analysis of the empirical data will be performed using Analysis of Variance (ANOVA). Essentially, this statistical

---

[17] http://irph-dev.upsi.edu.my/

procedure helps compare the mean scores of relevant variables among the three groups of the same population. The following is the information related to the experimental study to be carried out.

- *Sample:* 20 trainee teachers who will be undergoing teaching practicum at several selected schools.

- *Variables :* The independent variables are the criteria of the DLP that will be created. The dependent variables are the matched returned result based on the specified criteria.

- *Null Hypothesis*: The differences in the mean scores of information quality among the three groups are not significant.

$$(H_0: \mu_1 = \mu_2 = \mu_3) \tag{1}$$

IV. CONCLUSION

Based on the comparison among existing ontology development methodologies, a new ontology development methodology was proposed for the lesson plan domain. In light of the discussed problems, this proposed methodology will serve as a comprehensive, systematic guideline to help system developers produce an ontology for other domains based on a knowledge representation that supports web semantics. The activities involved within each phase, and techniques applied for each activity were clearly explained. Ultimately, this guideline can help in the development of high quality ontology to support all users to perform their task with greater efficacy.

This methodology to support based on a web semantic application can help users gain access to information that is not only ample but also relevant to the preparation of lesson plan. This application can also help overcome the unmanageable amount of information typically produced with the use of normal search engines.

In addition, the same application can help overcome the limitations of databases based on attribute value knowledge representation, which invariably need constant updating by the system administrator. However, the terms in this lesson plan domain were mainly defined using Malay language which might limits the search result in other language. This can be overcome by using alternative terms in an international language such as English and Arabic which is a future plan for this research to support all users to perform their task with greater efficacy.

ACKNOWLEDGMENT

REFERENCES

[1] A. Dix, "Human-computer interaction," Springer, US, 2009, pp. 1327-1331.

[2] A. Newell, "The knowledge level," Artificial intelligence, 18(1), 1982, pp. 87-127.

[3] A. Öhgren, and K. Sandkuhl, "Towards a methodology for ontology development in small and medium-sized enterprises," In IADIS AC, 2005, pp. 369-376.

[4] A. Saad, "A case-based system for lesson plan construction," Doctoral Dissertation, Faculty Of Science of Loughborough University, 2011.

[5] A. Saad, P. W. Chung, and C. W. Dawson, "Effectiveness of a case-based system in lesson planning," Journal of Computer Assisted Learning, 30(5), 2014, pp. 408-424.

[6] C. Ramirez, and B. Valdes, "A General Knowledge Representation Model of Concepts," Advances in Knowledge Representation, InTech, 2012.

[7] Cambridge Semantic, 2016. Retrieved from http://www.cambridgesemantics.com/semantic-university/rdfs-vs-owl.

[8] D. H. Le, and V. T. Dang, "Ontology-based disease similarity network for disease gene prediction," Vietnam Journal of Computer Science, 2016, pp.1-9.

[9] G. Tokdemdr, "An Assessment Model For Web-Based Information System Effectiveness," Doctoral Dissertation, School of Informatics Of The Middle East Technical University, 2009.

[10] J. Angele, D. Fensel, D. Landes and S. Studer, "Developing knowledge-based systems with MIKE," Automated Software Engineering, 5(4), 1998, pp. 389-418.

[11] J. Mishra, and A. Mohanty, Software Engineering. New Delhi, India: Dorling Kindersley, 2012.

[12] M. Fernández-López, A. Gómez-Pérez, and N. Juristo, "Methontology: from ontological art towards ontological engineering," 1997.

[13] M. L. Brodie, and J. Mylopoulos, "Knowledge bases and databases: semantic vs. computational theories of information," New Directions for Database Systems, Ablex, 1986, pp. 186-218.

[14] M. Uschold, and M. King, "Towards a methodology for building ontologies," Edinburgh: Artificial Intelligence Applications Institute, University of Edinburgh, 1995, pp. 15-30.

[15] N. Xue, S. Jia, J. Hao, and Q. Wang, "Scientific ontology construction based on interval valued fuzzy theory under Web 2.0," Journal of Software, 8(8), 2013, pp.1835-1842.

[16] P. Hitzler, M. Krötzsch, and S. Rudolph, "Foundations of semantic web technologies," Chapman and Hall/CRC, Taylor and Francis Group, 6000, 2010, pp. 33487-2742.

[17] R. R. Starr, and J. M. P. De Oliveira, "Concept maps as the first step in an ontology construction method," Information systems, 38(5), 2013, pp.771-783.

[18] R. Subhashini, and J. Akilandeswari, "A survey on ontology construction methodologies," International Journal of Enterprise Computing and Business Systems, 1(1), 2011, pp.60-72.

[19] T. Aaberge, and R. Akerkar, "Ontology and Ontology Construction: Background and Practices," IJCSA, 9(2), 2012, pp.32-41.

[20] T. R. Gruber, "A translation approach to portable ontology specifications," Knowledge acquisition, 5(2), 1993, pp. 199-220.

[21] W. H. DeLone, and E. R. McLean, "The DeLone and McLean Model of Information Systems Success: A TenYear Update," Journal of Management Information Systems, 19 (4), 2003, pp. 9-30.

[22] Y. Sun, and Z. Li, "Ontology-based domain knowledge representation," In 2009 4th International Conference on Computer Science and Education, 2009, pp. 174-177.