Original Research Paper

Development of Domain Ontology for a Semantic Online Forum

Hazalina Hashim and Shahrul Azman Mohd Noah

Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, Selangor, Malaysia

Abstract: Ontology has the potential to structure domain knowledge and formally represent the knowledge in a more precise, explicit, consensus-based and meaningful manner. However, to our knowledge, none have made an effort to integrate forum discussion with subject module knowledge into ontologies to enhance learning of the subject taught. As such, it is the interest of this paper to reuse, classify, integrate and semantically organize both of this knowledge by means of ontological structure. Thus, detail of the ontology development process is presented consisting of two main phases: Namely knowledge acquisition (from both the subject module and forum discussion) and knowledge modeling. The resulting ontology was evaluated via the use of a prototype, a concept annotation task and questionnaires. Findings indicate that the proposed ontology model is sufficient to accommodate learning needs according to the assessment of experts.

Keywords: Domain Ontology, Ontology Development, Semantic Web, Forum Discussion

Introduction

Studer *et al*. (1998) define ontology as "knowledge that being structured in the sense of taxonomies and reflect some consensus". Ontology is used to formally represent the knowledge structure of learning materials (Chi, 2009) and share the domain knowledge through modeling and the creation of concepts and relationships between those concepts (Noy and McGuinness, 2001). The use of concepts and its relationships is similar to a dictionary or glossary, but with richer structure, relationships and axioms that describe the domain knowledge more precisely (Millard *et al*., 2006).

One of the most widely used tools to further support learning in the e-learning education is online forum discussion (Macdonald, 2008). Over the time, forum discussion that use to resolve issues builds up a very large archive of question and answer knowledge repositories across subjects. Therefore, mining and reusing such resources is highly desirable and valuable to enhance learning of the subject taught. Although many researchers claim the importance of reusing such knowledge, few have took efforts to compile, organize and represent this implicit knowledge. This is because such knowledge is not formally organized and scattered throughout huge textual resources. Furthermore, classifying, extracting and accessing relevant knowledge

of interest from such a pool of "unstructured" question-answer resources are challenging tasks. What is required is a method to semantically organize the resources by means of ontology that are able to facilitate searching and reusing them.

Therefore, this study aims to reuse and structure the knowledge from subject module and forum discussion by means of ontological structure to enhance learning of the subject. The development applied the formal ontology engineering method and subsequently evaluated the ontology model through a prototype used, a concept annotation task and questionnaires.

Literature Review

Abel *et al*. (2010) highlight that the way forums discussion are currently designed, makes it difficult to classify and access relevant threads of interest. The drawbacks appear in several situations: (1) when learners start discussions in the wrong thread, (2) when learners post many different topics or ideas in a single thread, (3) when learners post the same topic in different threads, (4) when other learners respond to posts created in any of these situations themselves create further discussions or sub posts that are not in the correct thread and (5) when too many posts are made in a single thread.

Li *et al*. (2009) mentioned that simple keyword search and navigation implemented in most forum

Science Publications

discussion tend to be inadequate, resulting in a large number of irrelevant messages. In addition most users cannot adequately express their interests through keyword-based queries (Sieg *et al*., 2004). As a result, learners face the daunting problem of how to find what they need (Dicheva and Dichev, 2009). Studies also indicate that forum discussion filled with a large number of irrelevant posts with inaccurate and personal information (Hough *et al*., 2004; Stahl *et al*., 2006). This situation has made the forum a noisy environment where learners difficult to focus on learning (Wang *et al*., 2010) but more on managing their personal information rather than using the platform for collaborative knowledge building (Humayoun *et al*., 2012).

The effort to reuse knowledge from forum discussion, has been ignored except in the work of Li *et al*. (2009). Li *et al*. (2009) have designed an enhanced semantic forum system, based on the w3china discussion board as a data source that focuses on the semantic organization and association of discussion transcripts based on a domain ontology and text mining technologies. The system offers semantic searching with relational navigation support to guide learners through well-structured and coherent messages relevant to their learning needs. Our work is quite similar to (Li *et al*., 2009), as we also design a semantic forum system based on domain ontology modeled. However, our data source is not limited to forum discussion only; textbooks are also sourced for semantic organization and association. In contrast with the work of Li *et al*. (2009) our study focuses on developing a tool to enhance learning by filtering and analyzing a source specific to the subject matter, as well as through discussions with subject matter experts, tutors and learners.

The use of ontology through semantic web technology implementation offer promising solutions to address these challenges. In addition, ontology has already been recognized and used in the e-learning education (Bedi *et al*., 2010; Chu *et al*., 2011; Pahl *et al*., 2010) with different purposes ranging from the definition of a domain-specific terminology to the use of conceptual models and inference in the generation and composition of learning content and systems. Specifically, there is a need for a system that can be dedicated to a community that defines its own vocabulary by means of an ontology and that can offer relevant resources at the appropriate time and deal with structural incoherence in order to meet learning demands (Abel, 2009; Li *et al*., 2009).

Ontology Development

The ontology development process consisting of two main phases: Namely knowledge acquisition (from both the subject module and forum discussion) and knowledge modeling. The development was based on the Uschold and Gruninger (1996) ontology engineering methodology, which consists of knowledge acquisition, coding and integrating existing ontologies. Further details for each phase are provided in the next section.

Knowledge Acquisition

The knowledge acquisition phase was referred to the Uschold and Gruninger method involving (1) identification of the key concepts and relationships in the domain of interest; (2) production of precise unambiguous text definitions for such concepts and relationships; (3) identification of terms to refer to such concepts and relationships and, finally, agreeing on all of the above. In this study, knowledge acquisition process refers to the act of collecting knowledge and capturing the domain of interest.

The semantic online forum in this study is concerned with question-answer knowledge from forum discussion and university's subject module that are associated with the ontological structures. Throughout this paper, the System Analysis and Design (SAD) subject is used to illustrate the process of domain ontology modeling and development to support semantic organization of forum discussion knowledge.

From our point of view, the knowledge acquisition process requires direct human involvement to clarify ambiguous terms, bilingual posting and the subjective nature of various metadata to better evaluate the meaning in the designed ontology. Noy and Hafner (1997) consider the knowledge acquisition process as one of the most difficult activities in ontology building, as it not only involves a subjective representation but also how such a representation is classified. In our work, this process was conducted by two domain experts. They are known as Subject Matter Experts (SMEs). The SMEs were selected based on five years of teaching experience on the SAD subject in face-to-face tutorial classes as well as in the online forum.

Knowledge Acquisition from the Subject Module

A subject module refers to a unit of teaching or an academic course taught in a semester or one academic term. They contain well-organized structured content on the subject of concern. The knowledge acquired from the subject module provides a good grounding for the subject, presented in a hierarchical structure with chapters, sections and subsections. Figure 1 illustrates the process involved in acquiring knowledge from a subject module.

The first process began by structuring the knowledge in concept categories to reflect the essential aspect or big picture of the subject domain. The concept was defined as key terms or keywords discussed in the subject module. The concept categories were defined to

represent the delivery sequence of the knowledge structure in the subject module. It is similar to the idea of a backbone taxonomy, which consists of all of the rigid properties in the ontology, organized according to their subsumption relationships and represents a view of the ontology showing all of the most important properties, i.e., those that cover the entire domain (Guarino and Welty, 2002). Ferndndez *et al.* (1997) emphasize grouping concepts in appropriate categories so that learners are able to see the entire structure of the subject to understand what is important and what should be learned. A top-down approach was used to present the structure of the SAD subject module in a big picture perspective before looking at the concept details. The SMEs were seated together to discuss the appropriate structure and, finally, to share common understandings of the seven concept categories suggested for the SAD subject, namely: *Information System, Planning, Analysis, Design, Implementation, Maintenance* and *Object-Oriented*.

The second process was to determine concept relationships. This process was referred to concept relationships used in Simple Knowledge Organization Systems (SKOS). SKOS functions to represent thesauri, classification schemes, taxonomies, subject-heading systems or any other type of structured controlled vocabulary. SKOS aids were used in the organization and design of the ontology model prior to implementation. Such considerations were also given by the studies of (Carneiro *et al.*, 2010; Gerbé and Kerhervé, 2010; Summers *et al.*, 2008) to manage knowledge in many domains using SKOS and present the knowledge in web applications. The concept relationships were also structured upon the aforementioned concept categories. Fig. 2 indicates the concept relationship types that were used in this study. Details on the concept relationships types are further explained in the knowledge modeling section.

The third process was concept identification using the middle-out and key terms approaches. The middle-out approach identified concepts in the subject module based on the semantic relationships *narrower, broader* and *related*, as in Fig. 2. In brief, *broader* relationships relate to a concept that is more general in meaning, *narrower* relationships relate to a concept that is more specific in meaning and *related* relationships refer to concepts that are associated with meanings. The key terms were referred to terms listed at the end of each topic in the SAD module, as in Fig. 3. In this case, terms such as '*business process re-engineering*,' '*functional requirements*' and '*interviews*' are potential terms to be considered as concepts. This approach helped to identify and gather all of the useful and potentially usable domain knowledge and its meanings.
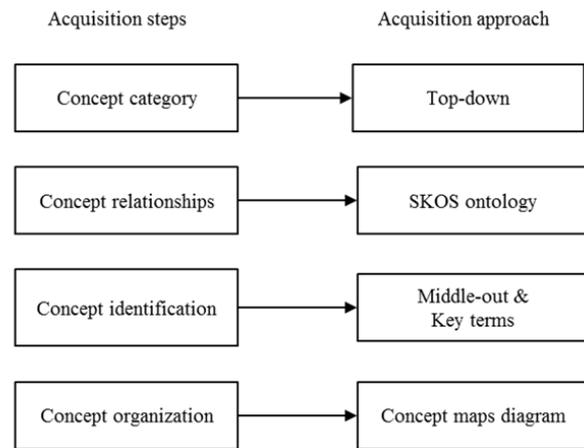

Fig. 1. The process of acquiring knowledge from a subject module
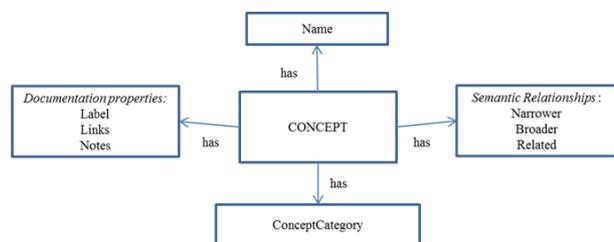
Fig. 2. Model of concept relationship types

| Key Terms	
Business process re-engineering (BPR)	Owners
Functional requirements	Questionnaires
Interviews	Structured Questions
Joint application design (JAD)	Stakeholders
Non-functional requirements	Technical staff
Observation	Users

Fig. 3. Sample of key terms

The fourth process was to organize the concepts identified using concept maps. Novak and Gowin (1984) define concept maps as "tools for graphical representations of knowledge that comprise of concepts and the relationships between them". Concept mapping can be seen as a first step in ontology building and can also be used flexibly to represent knowledge structure for meaningful learning (Chen *et al.*, 2008). The researcher assisted this process by drawing the concept identified from SMEs into concept map diagrams using an open-source mind mapping application, Free Mind. The use of concept maps, as shown in Fig. 4, was able to

facilitate verification by SMEs at the earliest possible stage to reduce ambiguous synonymous concepts as well as vague concept relationships. For example, in the SAD subject, *client server concept* was expressed ambiguously with ending terms such as *architecture, model, format, system*; *alternative design strategy* or *alternative solution concept* expressed synonymously as *design strategy concept*; *deployment concept* was expressed synonymously as *implementation concept*; the relation of *structure chart concept* was vaguely assigned under the relationship of *analysis concept* rather than *development tools concept*.

A total of 135 concepts were identified and subsequently linked to its relationships, as illustrated in Fig. 5. Differences that exist in concept map diagrams were compared, discussed and referred against the subject module for several times until both SMEs reached a final

agreement. The results of this   process from the subject module were captured in the subject ontology by the researcher in the knowledge modeling section.

Knowledge Acquisition from Forum Discussion

Forum discussion offers a list of discussion topics called threads. The threads are mainly organized by folders following the topics in the subject module. However, there are some other common folders created by the forum administrator, such as Introduction, General Concerns, Discussion on Assignment and Discussion on Final Exam. Each forum thread usually contains an initiating post and a couple of reply posts. The initiating post usually contains several questions and reply posts may contain answers to the questions in the initiating post or new questions. It is possible or even common for learners to pursue different questions in parallel within one thread.
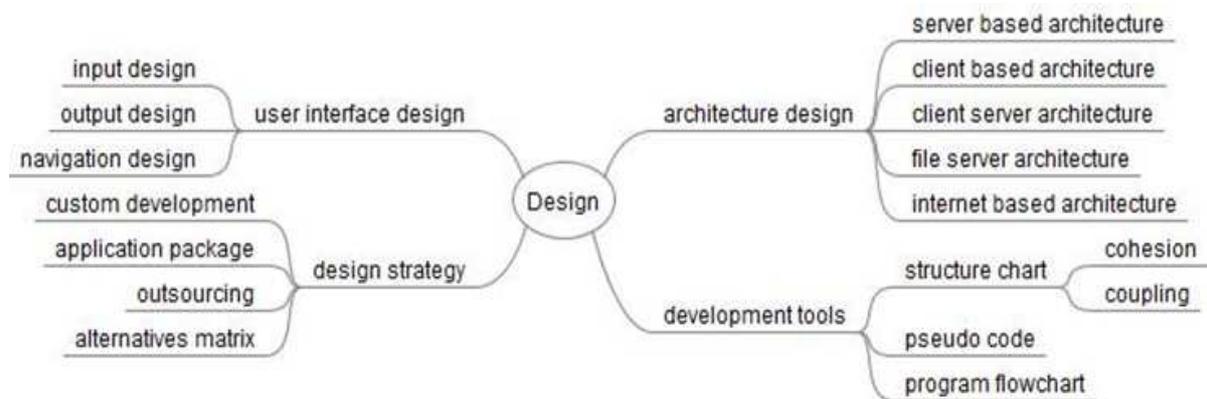
Fig. 4. Example of a concept map diagram
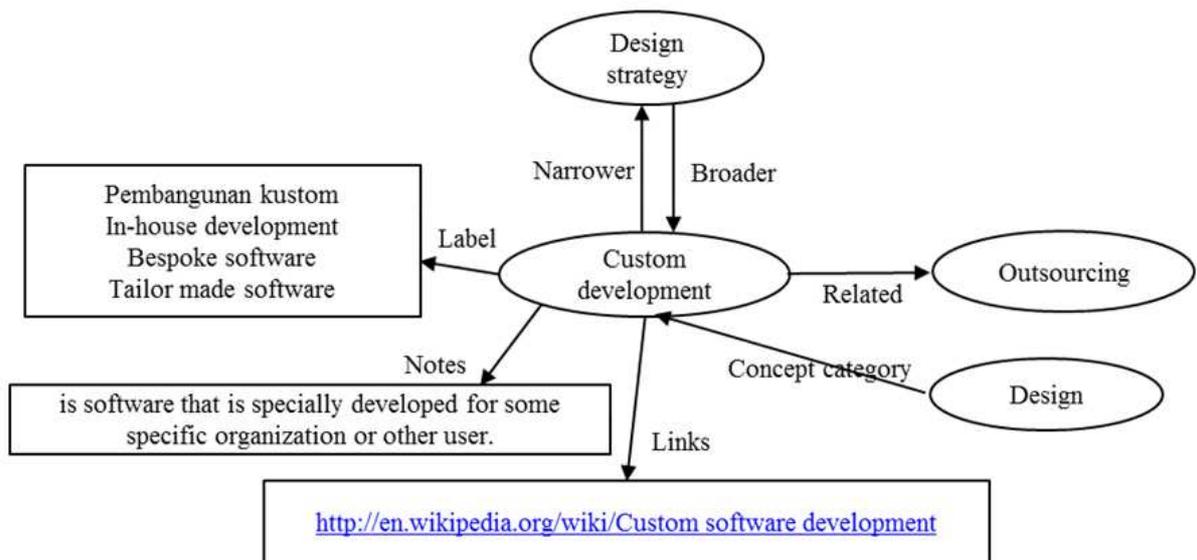
Fig. 5. Example of a concept map and its relationship to other concepts and details

Thread posting can be inclusive of announcements, greetings, chatting, questions or answers. Due to the existence of different types of information, not every thread in forum discussion is about questions or answers. The number of learners participating in the discussion also varies according to the number of learners registered for the subject in that particular semester, which may range from 55 up to 160 learners per discussion. The forum discussion is led by one online tutor, whose function is to actively guide the discussion with questions, advice, monitoring and simulating of interaction.

All discussions between learners and tutors in the online forum were stored in the discussion transcripts. These transcripts can be used by learners for reflective purposes, or they can serve as data for research (Meyer, 2004). The transcript is valuable in that it provides an accurate record of nearly all of the dialogue and interaction that took place. The transcripts in this study were taken from the SAD subject forum discussion in the university's Learning Management System (archive section) for the years 2009 through 2012, collecting 15 transcripts in total. The same SMEs as in the previous process were asked to identify, from the printed transcripts, what parts of the content represent questions and the corresponding answers. In this study, several steps were used in the acquisition process, as illustrated in Fig. 6, to assist the SMEs' judgment during the analysis of the discussion transcripts.

The first process was to identify questions in the transcript based on several approaches suggested by (Graesser and Person, 1994; Hong and Davison, 2009; van Zee and Minstrell, 1997). The process was identified based on: (i) interrogative expressions using what, where, when, why, who and how words; (ii) utterances that would be followed by a question mark; (iii) the position of the question, such as in the title of the thread; and (iv) authorship, which concerns who would usually ask questions in the discussion.

The second process was to identify answers in the transcript based on approaches suggested by (Harabagiu et al., 2001; Hong and Davison, 2009). The process was identified by the following: (i) the position of the answer post, which usually appears close to the question; (ii) understanding all replies to the question posted in order to gain a general idea of which reply post is an answer, as it would be less meaningful and more inaccurate to only extract the first reply as the answer; (iii) finding some of the words used in the question in the candidate answers. However, the words used in the question can be phrased differently or in different languages but still imply the same answer. Thus, synonyms and related terms that have similar meaning were used to identify answer candidates.

During these two processes, SMEs raised a concern for similar questions or answers identified in the transcript. Such similarities can be overcome by using various similarity measures, such as cosine similarity with the TFIDF weighting function (Li et al., 2009) and KL-divergence (Metzler et al., 2007). However, Jeon et al. (2005) state that the cosine similarity measure works poorly because it favors short answers. The shorter the answer is, the higher the cosine similarity scores are. Answers in the SAD transcript were usually short, but some can be very long because some tutors or learners generate answers by copying related answers from the web. Furthermore, Wang et al. (2010) mentioned that computing methods based on word co-occurrence are not effective to model question and answer similarities. Therefore, in this study, the SMEs were instructed to conduct the identification process without considering the similarity issues.
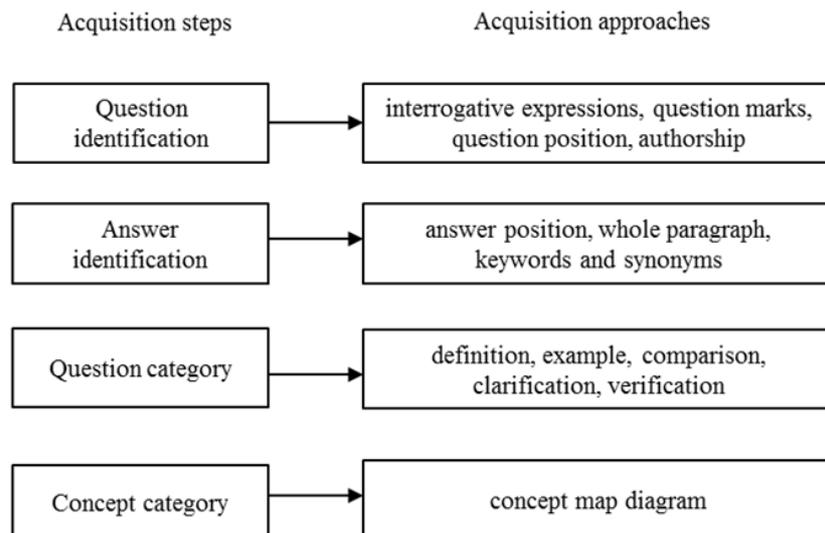
Fig. 6. Process of knowledge acquisition from discussion transcript
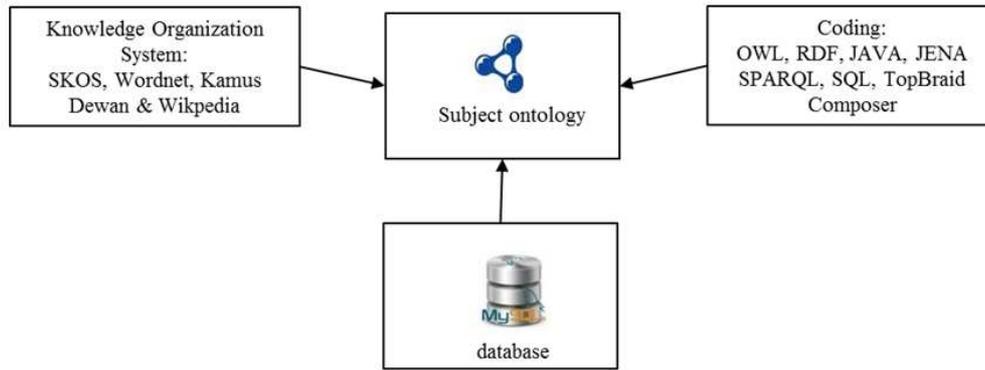
Fig. 7. Knowledge modeling for subject ontology
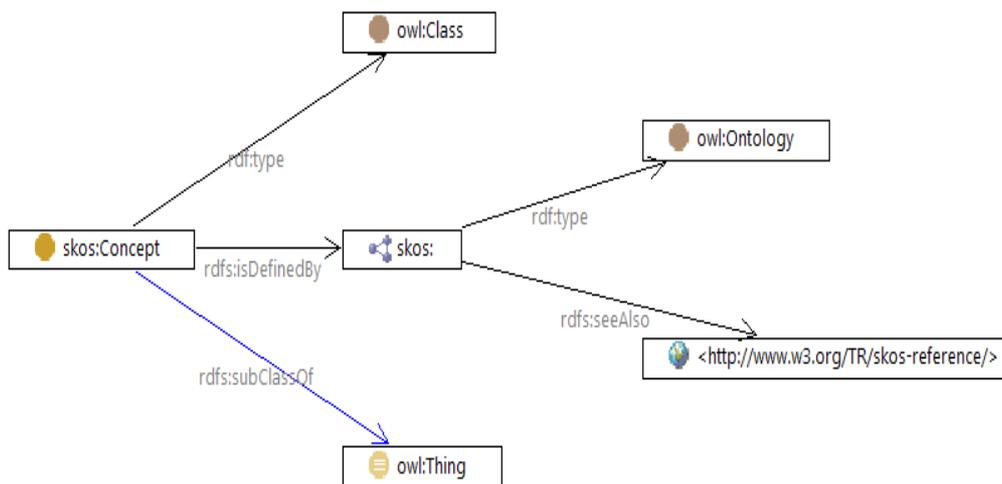
Fig. 8. SKOS as an RDF graph
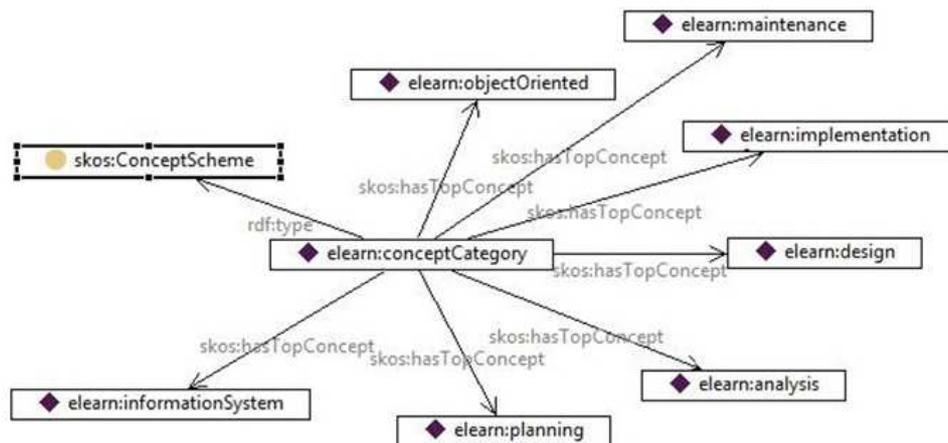
Fig. 9. Concept category model in the RDF graph

The third process was to categorize the identified questions into five question categories according to the (Graesser and Person, 1994) scheme. Table 1 shows the question category scheme with descriptions, indicators and examples. A set of indicators in English and Malay languages as an official language were used as a guideline to classify the questions into appropriate categories, as the transcripts contain both languages.

Finally, the SMEs were asked to classify the questions identified according to the concepts and concept categories they belonged to. To ease this process, concept map diagrams were used as well. As a result from this process, a total of 3219 messages were analyzed from 15 discussion transcripts which identified 320 questions and 543 answers.

Knowledge Modeling

The knowledge modeling phase models the ontology using the result of the knowledge acquisition process. The Uschold and Gruninger (1996) method suggests two processes in this phase, which are: (i) to reuse or integrate existing ontology or knowledge organization systems; and (ii) to encode the ontology in a formal and expressive representation language. In this study, three components were involved in modeling the subject ontology, as shown in Fig. 7.

Knowledge Organization System

To support the effort of standardization and integration for future use, this study extends existing World Wide Web Consortium (W3C) standard ontologies, namely Simple Knowledge Organization Systems (SKOS). It is recommended to consider existing ontologies in the same or similar domain or purposes when building the ontology to maximize the reuse of semantics (Noy and McGuinness, 2001; Uschold and Gruninger, 1996). SKOS is comprised of a set of RDFS classes and RDF properties that can be used to express the content and structure of a concept scheme, as illustrated in Fig. 8.

Table 2 describes some of the classes (beginning with an upper-case letter) and properties (beginning with a lower-case letter) defined in the SKOS ontology used in this study.

SKOS defines the concept category relationship using the *skos: ConceptScheme* class and the *skos:hasTopConcept* property, as shown in Fig. 9.

SKOS defines concept relationship types using *semantic relationships* properties by means of *broader*, *narrower* and *related*. Fig. 10 illustrates an example of

semantic relationships among instances and concepts. In this case, *structure chart* was a type of concept that has a *broader* concept named *development Tools*, *narrower* concepts named *cohesion* and *coupling* and *related* concepts named *pseudo Code* and *program Flowchart*. Organization by *semantic relationships* properties provided better control of the level of detail and was best for reflecting content covered in the subject module.

SKOS uses *documentation* properties to provide descriptions and additional information for a concept. Word Net (Miller, 1995) and Kamus Dewan (Baharom, 1994) were referred for *skos:prefLabel* or *skos:altLabel* properties to capture abbreviations, acronyms, synonyms, near-synonyms and alternative words to be represented by such a concept. Lexical resources, such as Word Net, can be very helpful in capturing the nuances of the English language, providing both generality and consistency (Guarino, 1997). Kamus Dewan is a Malay language dictionary published by Dewan Bahasa dan Pustaka, a government body responsible for coordinating the use of the Malay language in Malaysia. It is a useful reference tool for Malay literature as it provides suitable synonyms and abbreviations; it has also been used in computational linguistic research (Noah *et al.*, 2007). These properties give an advantage to the learner with regards to being able to recognize a concept even if they know the concept by a different name, as shown by example in Table 3. SKOS uses the *skos:links* property to refer to Wikipedia resources and the *skos:note* property to describe notes taken from a subject module of such a concept.

Coding

To be usable in the semantic context, with data in a well-defined and meaningful structure, the results generated by the knowledge acquisition on the subject module and knowledge modeling were compiled and mapped to the subject ontology in the Web Ontology Language (OWL). OWL is an RDF language developed by the W3C for defining classes and properties. OWL can be used to explicitly represent the meanings of concepts in vocabularies and the relationships between those concepts.
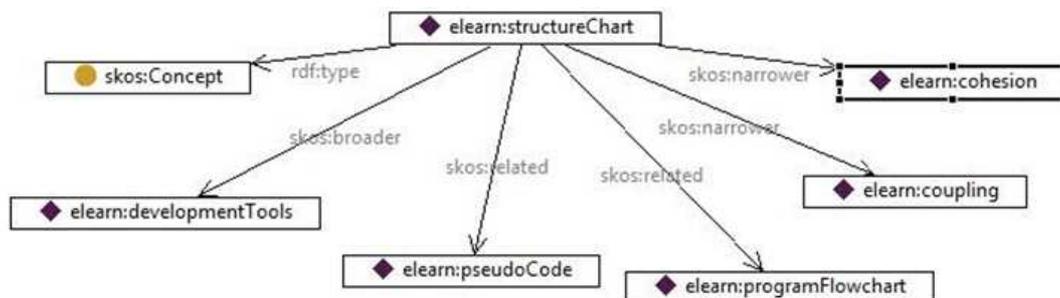

Fig. 10. Semantic relationships model in the RDF graph

Table 1. Question category scheme

Question category	Description	Indicator*	Example
Definition	Determine the meaning of a concept	Define terms… What is…	Can you define the functional requirement?
		Apakah itu… Takrifkan…	What is system testing?
Example	Request an instance of particular concept	State examples… Give example…	Can you give me an example of data modeling tools?
		Berikan contoh…	Can you state an example of decision tree?
Comparison	Establishing a comparison between two or more concepts	What is the difference between… Compare type of…	What are the differences between functional and non-functional requirements?
	Identify similarities and differences between two or more concepts	*Apakah perbezaan di antara…*	Compare the type of testing?
Clarification	Request an explanation, discussion or description of a concept	Explain, Elaborate, Describe, Discuss	What are the advantages of DFDs?
		How, Why, What are	Why does the system need maintenance?
	Setting out advantages or disadvantages of a concept Identifying causes and consequences of a concept	*Terangkan, Huraikan, apakah faktor*	
Verification	For yes/no responses to factual questions	Can, Is, It	Can it be considered an open question type, if the question is regarding user satisfaction?
	Clarifying understandings	*Adakah*	
		Bolehkah	Is it ok to assume requirements when developing a system?

Words in italics refer to the Malay language indicators

Table 2. SKOS class and properties

Term	Definition
skos:Concept	An abstract idea or notion; a unit of thought.
skos:ConceptScheme	A set of concepts, optionally including statements regarding semantic relationships between those concepts.
skos:hasTopConcept	A top level concept in the concept scheme.
Semantic relationships properties	
skos:broader	Relates a concept to another concept that is more general in meaning.
skos:narrower	Relates a concept to another concept that is more specific in meaning.
skos:related	Relates a concept to another concept with which there is an associative semantic relationship.
Documentation properties	
skos:prefLabel	The preferred lexical label for a resource, English language
skos:altLabel	An alternative lexical label for a resource, Malay language
skos:note	A general note for any purpose
skos:links	A general link for any purpose

Table 3. Examples of concept documentation properties

Concept	Label
Planning	System planning, *Perancangan*, Initial investigation
Feasibility	Feasibility study, *Kebolehlaksanaan*
Economic feasibility	*Kebolehlaksanaan ekonomi, Kesauran ekonomi*
Technical feasibility	*Kebolehlaksanaan teknikal, Kesauran teknikal*
Operational feasibility	*Kebolehlaksanaan pengeoperasian, Kesauran pengendalian*
Intangible benefit	*Faedah tidak ketara*
Tangible benefit	*Faedah ketara*
Preliminary investigation	*Penyelidikan awal, Kajian awal*
System request	*Kehendak system*, System acquisition, *Pemerolehan sistem*
Request for proposal	RFP, *cadangan permintaan*

Words in italic are in the Malay language

The Top Braid Composer tool was used to encode the subject knowledge in OWL. A framework in the Java language called Jena API and the SPARQL Protocol and RDF Query Language (SPARQL) were used to extract and write the data into OWL in the form of *triples* (*subject, property, object*). Jena is an open source Semantic Web framework for Java that provides an API to extract data from and write to RDF graphs. SPARQL is an RDF query language that is able to retrieve and manipulate data stored in RDF format. Apart from that, the results from the discussion transcripts and a user profile were captured in a database. This user profile contains details of the user, such as username, password, full name and role (student or tutor). This knowledge was then mapped to the subject ontology for semantic representation. Direct mapping were performed from database to OWL using the Java Database Connectivity (JDBC) connector API and Structured Query Language (SQL) queries were used to extract data from the database.

Ontology Evaluation

Having designed and constructed the subject ontology, the next step was to proceed with ontology evaluation. The methods of ontology evaluation used in this study referred to (Pattuelli, 2011) study, which were via the development and use of a prototype, a concept annotation task and questionnaires. The goal was to evaluate the quality of the ontology model by confirming whether the domain knowledge is adequate to represent the ontology according to the assessment of experts. Furthermore, these evaluation methods were also considered by (Almeida, 2009; Boyce and Pahl, 2007; Bright *et al*., 2012; Chu *et al*., 2011; Tankeleviciene and Damasevicius, 2009), who have made the same initiative to evaluate the quality of ontologies content by expert assessment and assess the adequacy of the ontology for its intended tasks and how well it represents the domain of interest. The evaluation setting in this study was conducted in the university's lab and was performed individually by three tutors as the experts, recruited through direct solicitation. Participants were selected based on five years of tutoring experience on the SAD subject in face-to-face tutorial classes as well as in the forum discussion.

Before evaluation takes place, participants were introduced to the semantic online forum prototype accessed through the Chrome web browser. Participant were asked to view concepts and relationship types presented through the prototype by maneuvering to concept map interfaces as in Fig. 11 and concept relationship interfaces as in Fig. 12.

Then, participants were asked to perform a concept annotation task based on a set of questions identified from the discussion transcripts. This was in line with the suggestion of Gruninger and Fox (1995) to use a list of questions that knowledge based on the ontology should be able to answer. Table 4 indicates the questions used, which were randomly chosen from the discussion transcripts to represent each of the topics from the SAD module.

In this task, specifically, participants were asked to perform a self-directed walkthrough on all of the concept map interfaces in the prototype. Along with the questions prepared, participants had to indicate their search process by circling the target concepts and numbering them according to the questions on the printed concept map diagram. They were also invited to annotate the concept maps and write down any questions, concerns or suggestions they might have.

Upon completion of the concept annotation task, participants were asked to answer the open questions in the questionnaire. The questionnaire, shown in Table 5, extended from evaluation criteria of (Gómez-Pérez, 2004; Vrandečić, 2009). The questionnaire included three criteria to evaluate the quality of the ontology model, namely (i) *coverage* to measures the extent to which the ontology model represents the subject domain, (ii) *structure* to measure the extent to which the ontology model the domain concepts and relationships adequately and (iii) *vocabulary* to measure the extent to which the ontology model had appropriate terminology or terms to represent these concepts. Furthermore, participants were also asked for their perception on the *usefulness* to measure the extent to which the ontology able to facilitate tutors to formulate learning needs via concepts and relationships as well as suggestions for the ontology modeled.

The concept annotation task revealed that participants were able to find the target concepts from the questions prepared. The results indicated that participant #1 circled 15 concepts, participant #2 circled 20 concepts and participant #3 circled 16 concepts. Participant #2 indicated as the highest concept circled; this can be explained, as the participant not only circled the target concepts but also the broader concepts. Participant #3 raised a concern regarding uncertainty of a suitable target concept for Q10; however, for this task, the participant circled "*Implementation*" as the target concept, while other participants circled "*testing*" as the target concept. Such differences were reasonable, as the concepts circled represented the question asked and the choice of concepts vary due to subjective interpretations. Such a difference was also reported by Noy and Hafner (1997) because conceptualization activities involve subjective representation. Thus, people may see and categorize things differently in their minds.

In the responses to the questionnaire results, all participants reported positively on the quality of the ontology model.

Fig. 11. Example of concept map interfaces in the prototype

Fig. 12. Concept relationship interface in the prototype

Table 4. Questions for the concept annotation task

#	Questions
Q1	What are the factors that influence the project request?
Q2	How can we assess the feasibility of a project?
Q3	What are the differences between DFD logical and DFD physical?
Q4	Explain the advantages and disadvantages of the in-house strategy.
Q5	Can you explain more about preventive and perfective maintenance?
Q6	What is the difference between an object and a class?
Q7	Madam, what is the approach that we can use to increase the speed of system development?
Q8	Can we assume our requirement? If not, do you have other methods?
Q9	Describe the differences between tangible and intangible benefits.
Q10	Why do we need to perform evaluation?

Table 5. Questionnaire to evaluate the quality of the ontology model

#	Questionnaire
1.	Are the concepts within the scope of the subject?
2.	Is the number of concepts well extracted from the subject?
3.	Are there any concepts you could not find and would have liked to?
4.	Are the levels of concepts sufficient?
5.	Is the concept categories structure appropriate?
6.	Is the organization of the concept clear?
7.	Are the concept relationships used adequately?
8.	Would you connect the concepts in a different way? If so, what would they represent?
9.	Are the terms used to represent the concepts clear?
10.	Are the label properties used rich enough to represent the concept?
11.	Would you use different words to express these concepts?
12.	Your perception on the usefulness of the ontology modeled?
13.	Your suggestions for improvement?

Participant positively responded on the *coverage* of the ontology model:

> *"...yes within the scope.", "...concepts well extracted.", "...concepts levels are adequate.", "...concepts are covered but need to update from time to time."*

Four additional concepts suggested by participants, written during the concept annotation task, are as follows:

> *"description"* concept for each of the concept categories,
> *"lifecycle"* concept to be added under the "development approach" concept
> *"in-house"* concept to be added under the "design strategy" concept
> *"evaluation"* concept to be added under the "implementation" concept

Participants positively responded to the *structure* of the ontology model:

> *"...appropriate concept categories and concept approach.", "...like the structure design to links among concepts.", "...the organization laid out nicely and adequate relationship proposed."*

Participants positively highlighted the *vocabulary* of the ontology model:

> *"…terms used are clear and rich.", "...bilingual vocabs are necessary.", "...reduce the effort on vocab searching on net, however it need to be maintained.", "... the vocab represent the subject taught."*

Participants' perceptions on the *usefulness* of the prototype and their *suggestions* for improvement are as follows:

> *"....the prototype is worth the effort, it's good to implement this idea to other subject as well."*
> *"...good approach in organizing the information, hopefully can produce a lot of searching features from this."*
> *"...really like the idea of having a link to Wikipedia and notes from module, hope to see more links attached."*
> *"....ontology is useful for fast reference, searching and save learning time."*
> *"...able to know what is available and what to be learn by the proposed model."*
> *"...easy to visualize subject structure through concept maps diagram."*

Discussion

The nature of forum discussion, which is largely unstructured and interpretative, posed a major challenge when attempting to classify, extract and access the relevant knowledge of interest. Furthermore, question and answer from the forum discussion that is used to resolve difficulties in learning is not being reused to further enhance the learning process. Currently, few have made efforts to compile, organize and represent the forum discussion knowledge. To our knowledge, none have made an effort to integrate forum discussion with subject module knowledge into ontologies. However, this effort is important to address, as (Guzdial and Turns, 2000) state that effective knowledge is that which is related and relevant to course learning objectives. Therefore, to accommodate the learning needs, it is vital to semantically organize such knowledge by means of ontology that is able to facilitate searching, organization and the provision of meaningful discussion content that is related to the subject taught.

In this study, the ontology development was based on the Uschold and Gruninger ontology engineering methodology, which mainly involves knowledge acquisition (both the subject module and forum discussion) and knowledge modeling. After development, ontology evaluation was conducted to evaluate the quality of the ontology model. The rationale for the evaluation was based on the assumption that a sound ontology model is the basis for an effective ontology implementation. The findings from the concept annotation task demonstrated that the ontology model was able to assist the search process by finding and suggesting appropriate concepts. Meanwhile, the questionnaire results revealed tutors' positive perception that the ontology model was adequate to represent the subject domain. Additionally, the tutors also found that the ontology model was useful in giving them some clues as to what was available and searchable. The building of ontology by making use of written documents (i.e., the subject module) and subject domain experts as implemented in this study was shown to produce a more adequate ontology model that meets the needs of the users. Such a finding was also reported by (Pattuelli, 2011), which stated that expert decisions do contribute in modeling a better ontology.

Conclusion and Future Works

This study presents a detailed ontology development process to semantically represent the content of the subject module and forum discussion in the form of ontology. Such an ontological representation is important within e-learning communities, particularly to support reusability of previous forum discussion and

to sustain the production of e-learning resources tailored to learners' needs. The result of knowledge acquisition and the modeling process produced subject domain ontology, with dependencies on concepts and relationships, as well as related questions and answers. Given the difficulties in this process regarding subjective interpretation and classification, as well as maintaining concepts and discussions that evolve over the time, this development is essential to reduce tutors' burden of posting the same content every semester. This instead promotes tutors' efforts in enhancing existing learning materials that are able to increase learners' understanding of the subject taught. In addition, this development also equips learners with structured learning material that is in line with the learning objectives and helps to enrich relevant discussions. This study contributes to the semantic-rich learning environment, as the integration of the subject module and forum discussion knowledge allows for the delivery of the course via a more innovative and productive learning system. This effort is important to address because e-learning education plays a vital role in building a connected and collaborative learning community.

A number of enhancements are possible by considering automated maintenance processes that are capable of tagging concepts from the subject module and forum discussion to the ontology automatically in order to maintain up-to-date and rich learning material. The ontology model also needs to be implemented in other subjects as well as with a large number of concepts to test the system's capability and transferability between subjects. As the ontology model was able to guide the delivery of the system, future efforts can deliver the system in a way that matches the preferred learning style of the user by varying the sequentialization of content elements.

Funding Information

Author's Contributions

Hazalina Hashim: Conception, design, acquisition of data, analysis and interpretation of data.

Shahrul Azman Mohd Noah: Reviewing it critically and give final approval of the version to be submitted.

Ethics

This article is original and contains unpublished material. The corresponding author confirms that all of the other authors have read and approved the manuscript and no ethical issues involved.

References

Abel, F., I.I. Bittencourt, E. Costa, N. Henze and D. Krause *et al*., 2010. Recommendations in online discussion forums for e-learning systems. IEEE Trans. Learning Technol., 3: 165-176. DOI: 10.1109/TLT.2009.40

Abel, M.H., 2009. Using a semantic forum as learning support. In: Web 2.0: The Business Model, Lytras, M.D., E. Damiani and P.O. de Pablos, Springer US, ISBN-10: 978-0-387-85894-4, pp: 1-22.

Almeida, M.B., 2009. A proposal to evaluate ontology content. Applied Ontol., 4: 245-265. DOI: 10.3233/AO-2009-0070

Baharom, N., 1994. *Kamus Dewan*: Dewan Bahasa dan Pustaka.

Bedi, P., H. Banati and A. Thukral, 2010. Use of Ontology for Reusing Web Repositories for eLearning. In: Technological Developments in Networking, Education and Automation, Elleithy, K., T. Sobh, M. Iskander, V. Kapila, M.A. Karim and A. Mahmood (Eds.), Springer, Netherlands, pp: 97-101, DOI: 10.1007/978-90-481-9151-2_17

Boyce, S. and C. Pahl, 2007. Developing domain ontologies for course content. Educ. Technol. Society, 10: 275-288.

Bright, T.J., E. Yoko Furuya, G. J. Kuperman, J.J. Cimino and S. Bakken, 2012. Development and evaluation of an ontology for guiding appropriate antibiotic prescribing. J. Biomed. Inform., 45: 120-128. DOI: 10.1016/j.jbi.2011.10.001

Carneiro, L.C., C. Sousa and A.L. Soares, 2010. Integration of domain and social ontologies in a CMS based collaborative platform. Proceedings of the Confederated International Workshops and Posters: On the Move to Meaningful Internet Systems, Oct. 25-29, Springer, Greece, pp: 414-423. DOI: 10.1007/978-3-642-16961-8_62

Chen, N.S., Kinshuk, C.W. Wei and H.J. Chen, 2008. Mining e-Learning domain concept map from academic articles. Comput. Educ., 50: 1009-1021. DOI: 10.1016/j.compedu.2006.10.001

Chi, Y.L., 2009. Ontology-based curriculum content sequencing system with semantic rules. Expert Syst. Appli., 36: 7838-7847. DOI: 10.1016/j.eswa.2008.11.048

Chu, K.K., C.I. Lee and R.S. Tsai, 2011. Ontology technology to assist learners' navigation in the concept map learning system. Expert Syst. Appli., 38: 11293-11299. DOI: 10.1016/j.eswa.2011.02.178

Dicheva, D. and C. Dichev, 2009. Authoring and Exploring Learning Content: Share Content by Sharing Concepts. In: Semantic Web Technologies for E-learning, Dicheva, D., R. Mizoguchi and G.J. Greer (Eds.), IOS Press, Washington, DC, ISBN-10: 1607500620, pp: 24-43.

Ferndndez, M., A. Gomez-Perez and N. Juristo, 1997. Methontology: From ontological art towards ontological engineering. Proceedings of the AAAI97 Spring Symposium Series on Ontological Engineering, Mar. 24-26, American Asociation for Artificial Intelligence, USA, pp: 33-40.

Gerbé, O. and B. Kerhervé, 2010. A model-driven approach to SKOS implementation. Proceedings of the 5th International Conference on Internet and Web Applications and Services, May 9-15, IEEE Xplore, Barcelona, pp: 484-488. DOI: 10.1109/ICIW.2010.79

Gómez-Pérez, A., 2004. Ontology Evaluation. In: Handbook on Ontologies, Staab, S. and R. Studer (Eds.), Springer Berlin Heidelberg, ISBN-10: 978-3-662-11957-0 pp: 251-273.

Graesser, A.C. and N.K. Person, 1994. Question asking during tutoring. Am. Educ. Res. J., 31: 104-137. DOI: 10.3102/00028312031001104

Gruninger, M. and M. Fox, 1995. Methodology for the design and evaluation of ontologies. Proceedings of the Workshop on Basic Ontological Issues in Knowledge Sharing, (IKS' 95), Montreal, Canada, pp: 1-10.

Guarino, N. and C. Welty, 2002. Evaluating ontological decisions with OntoClean. Commun. ACM, 45: 61-65. DOI: 10.1145/503124.503150

Guarino, N., 1997. Understanding, building and using ontologies. Int. J. Human-Comput. Stud., 46: 293-310. DOI: 10.1006/ijhc.1996.0091

Guzdial, M. and J. Turns, 2000. Effective discussion through a computer-mediated anchored forum. J. Learning Sci., 9: 437-469.
DOI: 10.1207/s15327809jls0904_3

Harabagiu, S., D. Moldovan, M. Pasca, R. Mihalcea and M. Surdeanu *et al.*, 2001. The role of lexico-semantic feedback in open-domain textual question-answering. Proceedings of the 39th Annual Meeting on Association for Computational Linguistics, (ACL' 01), Association for Computational Linguistics, USA, pp: 282-289.
DOI: 10.3115/1073012.1073049

Hong, L. and B.D. Davison, 2009. A classification-based approach to question answering in discussion boards. Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval, Jul. 19-23, ACM, USA, pp: 171-178. DOI: 10.1145/1571941.1571973

Hough, B.W., M.W. Smithey and C.M. Evertson, 2004. Using computer-mediated communication to create virtual communities of practice for intern teachers. J. Technol. Teacher Educ., 12: 361-386.

Humayoun, S.R., A. Poggi, T. Catarci and A. Dix, 2012. Task-based user-system interaction. KI - Künstliche Intelligenz, 26: 141-149.
DOI: 10.1007/s13218-011-0166-z

Jeon, J., W.B. Croft and J.H. Lee, 2005. Finding semantically similar questions based on their answers. Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval, Aug. 15-19, ACM, Brazil, pp: 617-618. DOI: 10.1145/1076034.1076156

Li, Y., M. Dong and R. Huang, 2009. Toward a semantic forum for active collaborative learning. Educ. Technol. Society, 12: 71-86.

Macdonald, J., 2008. Blended learning and online tutoring. Br. J. Educ. Technol., 39: 1147-1147.
DOI: 10.1111/j.1467-8535.2008.00908_22.x

Metzler, D., S. Dumais and C. Meek, 2007. Similarity Measures for Short Segments of Text. In: Advances in Information Retrieval, Amati, G., C. Carpineto and G. Romano (Eds.), Springer Berlin Heidelberg, Italy, ISBN-10: 978-3-540-71496-5, pp: 16-27.

Meyer, K.A., 2004. Evaluating online discussions: Four different frames of analysis. J. Asynchronous Learning Netw., 8: 101-114.

Millard, D.E., F. Tao, K., Doody, A. Woukeu and H.C. Davis, 2006. The knowledge life cycle for e-learning. Int. J. Continuing Eng. Educ. Life Long Learning, 16: 110-121.
DOI: 10.1504/IJCEELL.2006.008921

Miller, G.A., 1995. WordNet: A lexical database for English. Commun. ACM, 38: 39-41.
DOI: 10.1145/219717.219748

Noah, S.A., A.Y. Amruddin and N. Omar, 2007. Semantic Similarity Measures for Malay Sentences. In: Asian Digital Libraries. Looking Back 10 Years and Forging New Frontiers, Goh, D.H.L., T.H. Cao, I.T. Sølvberg and E. Rasmussen (Eds.), Springer Berlin Heidelberg,
ISBN-10: 978-3-540-77094-7, pp: 117-126.

Novak, J.D. and D.B. Gowin, 1984. Learning How to Learn. 1st Edn., Cambridge University Press, New York, ISBN-10: 0521319269, pp: 199.

Noy, N.F. and C.D. Hafner, 1997. The state of the art in ontology design: A survey and comparative review. AI Magazine, 18: 53-74.

Noy, N.F. and D.L. McGuinness, 2001. Ontology development 101: A guide to creating your first ontology. Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880, Stanford University.

Pahl, C., M. Javed and Y. Abgaz, 2010. Utilising ontology-based modelling for learning content management. Proceedings of the EdMedia: World Conference on Educational Media and Technology, Jun. 29, Association for the Advancement of Computing in Education, Toronto, Canada, pp: 1274-1279.

Pattuelli, M.C., 2011. Modeling a domain ontology for cultural heritage resources: A user-centered approach. J. Am. Society Inform. Sci. Technol., 62: 314-342. DOI: 10.1002/asi.21453

Sieg, A., B. Mobasher and R. Burke, 2004. Inferring user's information context from user profiles and concept hierarchies. Classification, Clustering Data Min. Applic., 563-573.
DOI: 10.1007/978-3-642-17103-1_53

Stahl, G., T. Koschmann and D. Suthers, 2006. Computer-Supported Collaborative Learning: An Historical Perspective. In: Cambridge Handbook of the Learning Sciences, Sawyer, R.K. (Ed.), Cambridge University Press, New York, pp: 409-426.

Studer, R., V.R. Benjamins and D. Fensel, 1998. Knowledge engineering: Principles and methods. Data Knowledge Eng., 25: 161-197.
DOI: 10.1016/S0169-023X(97)00056-6

Summers, E., A. Isaac, C. Redding and D. Krech, 2008. LCSH, SKOS and linked data. Sci. Services Agents World Wide Web, 20: 35-49.
DOI: 10.1016/j.websem.2013.05.001

Tankeleviciene, L. and R. Damasevicius, 2009. Characteristics of domain ontologies for web based learning and their application for quality evaluation. Informatics Educ., 8: 131-152.

Uschold, M. and M. Gruninger, 1996. Ontologies: Principles, methods and applications. Knowledge Eng. Rev., 11: 93-136.
DOI: 10.1017/s0269888900007797

van Zee, E. and J. Minstrell, 1997. Using questioning to guide student thinking. J. Learning Sci., 6: 227-269.
DOI: 10.1207/s15327809jls0602_3

Vrandečić, D., 2009. Ontology Evaluation. In: Handbook on Ontologies, Staab, S. and R. Studer (Eds.), Springer Berlin Heidelberg,
ISBN-10: 978-3-540-92673-3, pp: 293-313.

Wang, B., X. Wang, C. Sun, B. Liu and L. Sun, 2010. Modeling semantic relevance for question-answer pairs in web social communities. Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics, Uppsala, Sweden, pp: 1230-1238.