New challenges for eLearning: supporting learning content through semantic web technology

Hazalina Hashim
Open University Malaysia
hazalina@oum.edu.my

Abstract

Nowadays, the web has become an important learning platform. Its accessibility has made it a successful environment particularly for eLearning education. However, the emerging technology of the Semantic Web sets out challenges that need to be explored in bringing the eLearning platform to the next level. This paper discusses the potential of semantic technology for supporting learning content in the eLearning environment. A conceptual model is referred to present the idea of learning content development through semantic web technology in the educational context. Furthermore, the ontological development model describes the process involved in building concepts and ontologies from learning content. Lastly, this paper also explains the impacts of semantic technology that contribute to the changes of a longer-term aim of eLearning education.

Introduction

Currently, the World Wide Web is primarily composed of documents written in HTML (Hyper Text Markup Language), a very useful language for visual presentation. HTML is a set of "markup" symbols contained in a Web page intended for display on a Web browser. Most of the information on the Web is designed only for human consumption. Humans can read and understand Web pages but their inherent meaning is not shown in a way that allows their interpretation by computers. One way to enable machine-to-machine exchange and automated processing is to provide the information in such a way that computers can understand it. This is precisely the objective of the semantic Web is to make possible the processing of Web information by computers.

Semantic Web is an extension of the current Web in which information is given well-defined meaning, better enabling computers and people to work in cooperation. (Berners-Lee, Hendler et al., 2001). The semantic web initiative aims to support explicit semantics and its automated processing (W3C, 2006a). E-learning is an area that can benefit from Semantic Web technologies. Recent advances in technologies for Web-based education provide learners with a broad variety of learning content available. Adaptive support based on learners' needs, background and other characteristics can help in selecting appropriate learning and during the learning process.

Numerous resources may be used during eLearning education. Some are internal and made by several factors implied in the eLearning, others are available on the Web, such as online courses, online forums, course supports, slides, bibliographies, frequently asked questions, lecture notes, and so forth. Ontologies are a way of representing such formal and shared information. They can be used to index data indicating their meaning, thereby making their semantics explicit and machine-accessible. Ontology technology as the knowledge representation and inference core of the Semantic Web, promises this

wide applicability (Berners-Lee, Hendler, & Lassila, 2001). An area such as education, where access to information is central, depends on the representation and organization of knowledge both for the content but also the meta-data level.

This paper represents research efforts to discuss the potential of semantic web technology for eLearning content development process. The discussion starts by referring existing model for content development, followed by ontological development process and finally the impact of semantic web technology in the eLearning education.

Conceptual Model of Learning Content Development

In this study, a referred conceptual model (based on Pahl & Holohan, 2009) shown in Figure1 below is used to adapt the potential of eLearning content development using semantic web technology for our eLearning system. This model provides an overview of content development process as a whole by defining the information exchanged (input, results, etc) and actors' role for each activity.

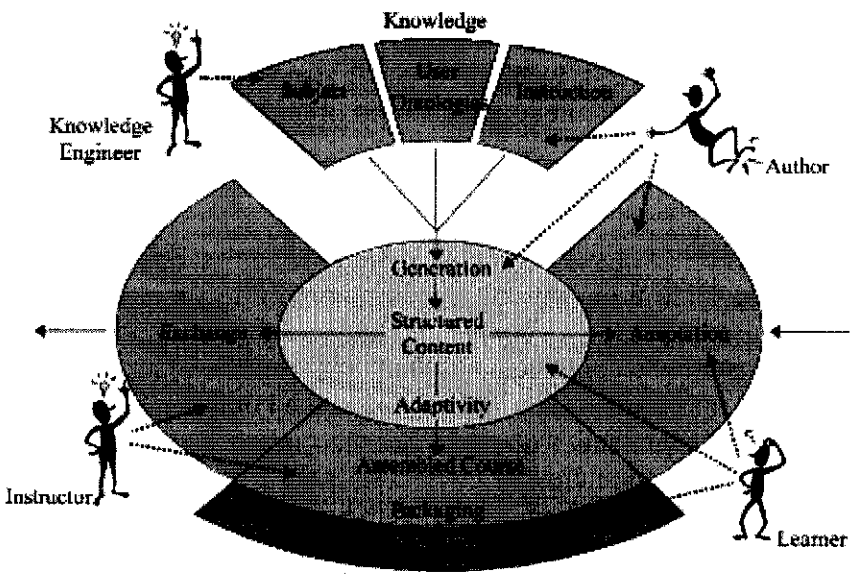

Figure 1: A conceptual model of learning content development

Based on the model in Figure 1, authors (Pahl & Holohan, 2009) have identified seven classifications of different forms of application using semantic web technologies for learning technology systems. These application types are organized into four contexts, determined by the activities and actors. The relevancy of the model referred explained in the next section from the perspective of our eLearning content development and actor's role.

The roles of the actors and activities involved are as follows:

- Knowledge engineers, or known as ontologists, are needed to develop subject, user and instruction ontologies. Subject domain ontologies are used to classify learning objects according to their subject. Instruction ontologies are used to represent the context of the subject which require education expertise as input and some activities

such as scenarios, links, examples, exercises, topic analysis or discussion. User ontologies describe the user type, the user's knowledge level and user preferences.

- The author, or known as subject matter expert, plays the crucial role in monitoring the overall development process. The focus is on the ontology-aware creation and organization of individual content units or known as learning object. It involves the creation and generation from unstructured content to structured content with content adaptivity. The task of content adaptivity to adapt the content to the needs or preferences of the user which requires the matching of learner's knowledge with knowledge represented in the learning content.
- The instructor, which can be a tutor or subject matter expert, assembles the learning object to larger units of study and packages it to particular subjects taught. Packaging needs to consider the interoperability, organization and sequencing of learning objects in an educationally sound way. A learning object is any chunk of learning material, regardless of whether it is a small piece or whole content. An important component of learning objects is learning object metadata (LOM) (Brase and Nejdl, 2003), or resource profiling (Downes, 2004) to facilitate standardisation of learning content.
- Once this eLearning content is ready by courses or subjects, instructors and learners can start their learning process by using the content available. Sharing and reuse are most important aspects for all actors involved. Learners, instructors and author can freely contribute to annotate the eLearning content from time to time. Instructors play an important role for exchanging content essentially to share content resources with the learners.

Ontology Development

Compared to traditional learning in which instructors play intermediate role between learners and learning content, the learning scenario in eLearning is completely different. Instructors no longer control the content delivery and learners have a possibility to combine learning content by their own. However, regardless of the time or expense put into creating advanced learning content, the content is useless unless it can be searched and indexed easily. This is especially true as the volume and types of learning content increase.

One solution to this problem is to use ontology. Ontology is defined as "an explicit specification of a shared conceptualization" Gruber (1995), where conceptualization refers to the objects, concepts, and other entities that are assumed to exist within some domain of interest and the relationships that hold among those entities. In other words, ontology helps us to make the knowledge represented in the eLearning content explicit. The significance for using ontologies are the sharing of common understanding between human and computer, reuse of domain knowledge, making domain assumptions explicit, the separation of domain knowledge and the operational and analysis of domain knowledge.

Ontologies have the potential to facilitate the creation of semantic relationships between various pieces of relevant and useful information, which is the backbone of semantic web, to enhance the learning experience in an eLearning environment. Ontologies can also facilitate provision of consistent vocabulary and word representation for clear communication within knowledge domain. Ontologies have been widely applied in the

context of integration and representation of various knowledge resources (Berners-Lee et al., 2001).

Before we can proceed with ontology development process, which plays a main component in the conceptual model explained in Figure 1, the concept and relationships understanding from learning content need to be addressed very clearly. Figure 2 explains the idea of general process of concept and its relationships (based on Boyce & Pahl, 2007). The advantage of separating concepts from learning content is that content aspects in the system can be changed without affecting the overall structure or vice versa. In fact, learning content (either learning objects or fragments) can address different learning aspects. Some might define and explain concepts and others might provide examples of a concept.

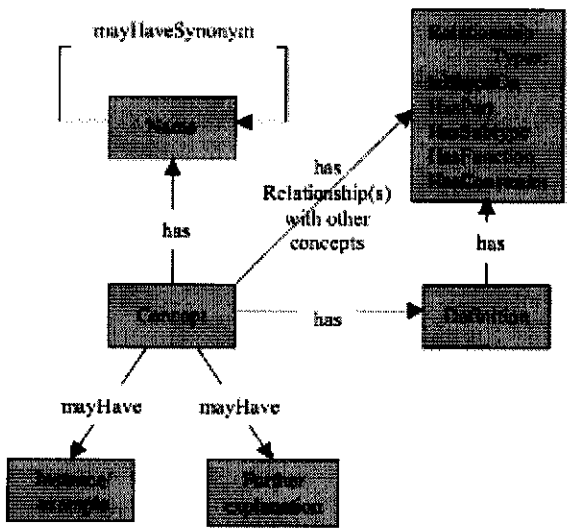

Figure 2: Overview of Concept Modeling and Relationships

Organizing concepts into hierarchy is a key aspect of building ontologies. The idea is similar to the process of abstraction and inheritance in object-oriented programming. The concept modelling explained in Figure 2 is also used as part of the ontological development process known as concept identification and concept organization. Figure 3 further explains the overall component involved in ontological development process.

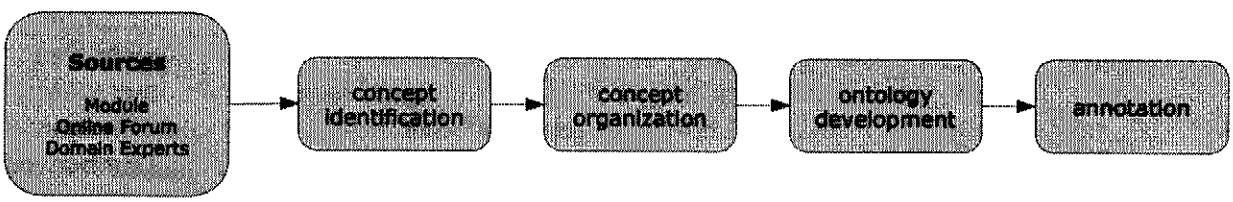

Figure 3: Overview of Ontological Development process

Initially the author starts by outlining desired subject content as set of concepts and associated relationships. This process is known as concept identification. The author can refer to existing resources such as modules, domain experts and online forum as source of the process. There are advantages of using each of the sources. For example, the module provides a good grounding of the subject and when each new topic is introduced, new terms are explained, thus providing the basis for concept identification.

Though not following the sequence, the online forum generally discussed various topics and using thorough observation authors can identify which learning concept is mostly being discussed and how comprehensive it should be.

- Domain experts can play their role in finalizing the concept identified and concept organization.
- The next task is to organise the concept in a top-down approach which means the most general concepts will be explained and evolve to more specific concepts. This structure matches the format provided in the module. Collection of learning concept will be used to organize into a larger learning concept based on inherent dependencies and sequencing that is derived from the subject content. In the scenario of content organization, knowledge is explicit and separated from the representation of content in the form of learning objects (Garlatti & Iksal, 2003). Different types of relationships have to be dealt with in the process of arranging the content units in a suitable sequence (Fischer, 2001).
- The next task is to develop ontologies by knowledge engineer using ontology tools provided. Subject ontologies can be richer than a vocabulary or concept hierarchy. Often it forms a conceptual model describing a full domain. The knowledge in instructional ontologies can comprise a vocabulary to classify educational units, such as definition, example, or exercise. User ontologies comprise user profiles such as learning styles, past experiences and motivation. These ontologies will use the sequencing algorithm in the ontological reasoning facilities of an underlying logic to determine the ordering dependencies.
- In order to support sharable resources from eLearning content that have gone through the ontological development process, finally, author, learners and instructors will involve in semantic annotation process. Ding (2005) defined semantic annotation as "...a process to label web page content explicitly, formally, and unambiguously using ontologies". In addition, Scerri et al. (2005) defined it as "...the process of attaching semantic descriptions to Web resources by linking them to a number of classes and properties defined in ontologies." Learning Object Metadata standard LOM (IEEE 2002) provides a metadata framework for the annotation of learning objects. LOM defines the attributes required to fully describe a learning object.

The main part of Semantic Web is that ontologies should provide a formal description for a shared domain conceptualization. Ontologies applied to the Web are creating the Semantic Web (Fensel et al, 2001). In practice, ontologies are often developed using integrated, graphical, ontology-authoring tools, such as Protégé (Protégé, 2009), OntoEdit (Sure et al., 2002), TM4L (Dicheva and Dichev, 2007) and HOZO (Hozo, 2009). They are used to develop new ontologies and modify existing ones. They allow the author and ontologist to edit and develop ontologies concentrating on the domain's concepts and relationships, without worrying much about ontology-representation languages. The author can choose ontologies from a list, choose attributes and relations from another list, edit, add, remove, and merge ontologies. The output is usually produced in a specific high-level ontology-representation language such as OWL, RDF/RDFS, HTML, or in plain text.

RDF (Resource Description Framework) and OWL (Web Ontology Language) standards enable the Web to be a global infrastructure for sharing both documents and data, which make searching and reusing information easier and more reliable as well.

RDF is a standard for creating descriptions of information, especially information available on the World Wide Web. What XML (Extensible Markup Language) is for syntax, RDF is for semantics. OWL provides a language for defining structured Web-based ontologies which allows a richer integration and interoperability of data among communities and domains.

Technology Benefits

The key ideas of semantic web, namely common shared meaning (ontologies) and machine-processable metadata, establish a promising approach for satisfying the eLearning requirements. Expectations for this technology contribute to longer-term aim of educational change which are (a) increase the effectiveness of education, (b) to increase the flexibility and accessibility of education, (c) increase the attractiveness of education and (d) to decrease the workload for staff (or more in general: to decrease the institutional costs). Thus the relevancy of this semantic web technology for eLearning education depends on how much it contributes in the accomplishment of these aims.

When we are able to represent eLearning courses in a semantic way, it opens the possibility to realize the needs of eLearning education:

- *Sharing and re-use of learning content* is one of the major objectives in the field of eLearning, more specifically learning objects (Littlejohn, 2003). This sharing and re-use is needed to make course development more efficient; however sharing is hard to do when the learning objects are not semantically represented.
- *Adaptation to individual learner characteristics* is highly desirable since no two learners have the same learning pre-requisites, skills, aptitudes or motivations. However, such adaptation can only be done realistically when the adaptation is wholly or at least partially automated. Otherwise it becomes too much work for the learners and tutors.
- *Learning content* dynamically changes constantly through learners and tutors interactions, input, experiences or new practices. The Semantic web enables the use of distributed knowledge provided in various forms, enabled by semantically annotation of content. Distributed nature of the Semantic Web enables continuous improvement of learning content. This enables an effective co-operative content management.
- *Nonlinear Access.* Knowledge can be accessed in any order the learners wishes, according to his/her interests and needs. This can be done by performing semantic querying for the suitable learning content. Learning content can be retrieved in the context of actual problems, as decided by the learners.

These benefits could result in more flexibility in the creation and management of eLearning content and, as a consequence, will reduce cost. The wealth information available in our eLearning content can be harnessed using semantic web technology with incorporating pedagogical theories and processes to provide a complete paradigm change towards more active-learning environment. Ultimately, the learners will benefit from these through extended availability.

Conclusion

The purpose of this study is to discuss the potential of semantic web technology to be applied in supporting eLearning content effectively. These discussions lead with an overview of model in learning content development using this technology in eLearning platform. Then the overview of ontological development process is described to reuse and sharing the wealth information available in the eLearning resources which can be semantically represented using concepts and ontologies.

This new approach to produce semantic eLearning content from existing resources enables the next generation of eLearning intake will be better than the current offerings. The next stage of eLearning will see diversification of content and models better fitted to particular learning needs and applications, instead of a one-size-fits-all approach. By focusing on machine representation and understanding the content infrastructure, eLearning can be realized as efficient learning in more productive and innovative learning organizations.

References

Berners-Lee, T., Hendler, J. and Lassila, O. (2001). *The semantic web*. Scientific American, 284 (5), 34 - 43.

Boyce, S., & Pahl, C. (2007). *Developing Domain Ontologies for Course Content*. Educational Technology & Society, 10 (3), 275-288.

Brase, J. and Nejdl, W. (2003). *Ontologies and metadata for e-learning*. Springer Verlag. 579-598.

Ding, Y. (2005). *Study of Design Issues on an Automated Semantic Annotation System*. AIS SIGSEMIS Bulletin 2: 45-51.

Fensel, D., van Harmelen, F., Horrocks, I., McGuinness, D.L., & Patel-Schneider, P.F. OIL (2001). *An Ontology Infrastructure for the Semantic Web*. IEEE Intelligent Systems 16 (2), 38-45.

Garlatti, S. & Iksal, S. (2003). *Declarative specifications for adaptive hypermedia based on a semantic web approach*. Proceedings of the User Modeling Conference, 81-85.

Gruber, T. (1995). *Towards principles for the design of ontologies used for knowledge sharing*. International Journal of Human Computer Studies, 43 907-928.

Dicheva, D. and Dichev, C. (2007). *Helping courseware authors to build ontologies: the case of tm4l*. IOS Press, p.77–84.

Downes, S. (2004). *Resource profiles*. Journal of Interactive Media in Education, 5.

Fischer, S. (2001). *Course and exercise sequencing using metadata in adaptive hypermedia learning systems*. Journal on Educational Resources in Computing (JERIC), 1(1).

HOZO (2009). Hozo ontology editor, Osaka University and Enegate Co. Ltd. Available at: http://hozo.jp. Accessed on: 03/03/2009.

Littlejohn, A. (2003). *Reusing Online Resources: A Sustainable Approach to eLearning*. London: Kogan Page.

IEEE Learning Technology Standards Committee LTSC (2002). IEEE P1484.12/D4.0 *Draft Standard for Learning Object Metadata (LOM)*. IEEE Computer Society.

Pahl, C., & Holohan, E. (2009). *Applications of Semantic Web Technology to Support Learning Content Development*, Interdisciplinary Journal of E-Learning and Learning Objects, 5, 1-25.

PROTÉGÉ (2009). The protégé ontology editor and knowledge acquisition system. Available at: http://protege.stanford.edu. Accessed on: 03/3/2009.

Sure, Y., Erdmann, M., Angele, J., Staab, S., Studer, R. and Wenke, D. (2002). *OntoEdit: Collaborative ontology development for the semantic web*. Springer-Verlag. p. 221–235.

Scerri, S., Abela, C. and Montebello, M. (2005). *SemantExplorer: A Semantic Web Browser*. IADIS International Conference WWW/Internet, Lisbon, Portugal.