

Managing the Semantic Aspects of Learning using the Knowledge Life Cycle

Feng Tao, David Millard, Hugh Davis, Arouna Woukeu
University of Southampton
{ft, dem, hcd, awl}@ecs.soton.ac.uk

Abstract

In this paper we examine the semantic aspects of learning from both pedagogical and technological points of view. We suggest that if semantics are to fulfil their potential in the learning domain then a paradigm shift in perspective is necessary from information based content delivery to knowledge based, context-aware collaborative learning services. We propose a semantics driven knowledge life cycle that characterises the key phases in managing semantics and knowledge, and show how this can be applied to the learning domain.

1 Introduction

As e-learning applications become more integrated and e-learning systems more distributed there is an increased need to manage their software and data components [1]. There is a trend in the distributed computing and middleware areas of computing towards Service-Oriented-Architectures (SOA), and in particular the Grid is evolving as an SOA for securely orchestrating and sharing stateful services and resources across distributed organisations [2].

Both Web and Grid service architectures have been applied to the e-learning domain [3, 4], the argument is that they are advantageous as they are modular and extensible and offer increased interoperability. While Grid services were originally conceived as a method of distributing high performance computation, they also offer benefits in distributed knowledge and information management, offering functionality that is essential for serious e-learning applications, such as security and statefulness.

The semantic aspects of learning content are the key to facilitating large scale collaboration of e-learning activities over service-oriented infrastructures. In order to use explicit and accurate semantics, a consensus in the domain at the conceptual level is necessary, so that computer and human participants can understand and communicate.

An ontology is the best vehicle in this context to formally hold a specification (of the conceptualisation) that can be shared within the community to describe semantics accurately and consistently. It explicitly defines the domain concepts and their relationships and is similar to a dictionary or glossary, but with richer structure, relationship and axioms that describe a domain of interest more precisely.

These rich semantics offer both teachers and learners new opportunities for locating and reusing resources [5,6,15]. But defining the correct semantics for a learning application is

difficult, and maintaining ontologies can be problematic (akin to managing the evolution of a complex graph).

We propose a knowledge life cycle for learning, to help define and maintain evolving semantics [14]. Our intention is not to develop a definitive ontology, or to promote a particular architecture, but to demonstrate how a semantic-driven knowledge life cycle model can be applied to the learning domain.

2 A Pedagogical View of Semantics

In this section we examine the affordances of semantics from a pedagogical point of view, in an effort to answer the question: what can semantics do for the domain of learning?

2.1 How Semantic Enrichment can improve learning

Increased semantics offer students a more effective view of their learning and enables new learning opportunities [5,15]. There are a number of the ways in which reasoning about semantics can improve learning opportunities.

- **Connecting Communities:** Services can put people in contact with other people who are experts or learners with similar interests.
- **Personalised Content:** Intelligent tutoring systems have for some time been delivering content that was personalised for the user, based on an understanding of their goals and previous knowledge.
- **Personalised Sequencing:** Adaptive Hypertext Systems provide pathways through materials by matching domain ontologies with evolving user models.
- **Adaptive Assessment:** Systems may choose questions for the learner at the boundary of their understanding, thus improving the efficiency of assessment and providing feedback that provides detail in critical areas.
- **Recommender Agents:** The system could recommend alternative resources based on user searching and studying patterns. In a formal setting, it could query the syllabus and timetable to recommend a plan of study.
- **Annotation Tools:** Users could annotate information themselves, providing useful information for others and allowing both readers and other services the opportunity to process the information in alternative ways.
- **Search Engines:** When resources have been semantically enriched then search engines can be much more powerful. Where services are semantically enriched search engines can choose suitable services to manage the query.

2.2 How Semantic Enrichment can improve the Management of learning

E-Learning practitioners often comment that they believe they spend as much time organizing materials as they spend on teaching and the production of materials. We believe that semantics may ease this problem by helping with:

- **Locating Materials:** Production of teaching materials is a time consuming task. The ability to locate and to re-use existing materials is a primary motivation for providing metadata for learning resources.
- **Student Management:** An understanding of the role of teachers, students, assessors etc. makes the production of services for assigning students to the correct classes, discussion groups, experimental teams etc. possible.
- **Timetable Management:** An important task for teachers of on-line tasks is the timing of events, such as the release of new materials, the closing date of an assessment, the time of a group session, etc. The events can be made to happen automatically when described in some language such as IMS Learning Design.
- **Record Keeping:** Record keeping and quality assurance can require teachers to spend much time ensuring that all the results are kept in the correct places such as institutional enterprise systems, as well as made available for QA purposes. This is an obvious target for automation by services that understand the goals.

Much of what has been described in this section is in effect suggesting the orchestration of services to achieve some goal. For example, an assessment system might call a service to handle some marks. This service might then ask an enterprise system service to store the marks in a database; it might call a service to annotate the student records with the new information, and then might call an email service to inform the students of the need to update their personal development plans accordingly. We believe that appropriate semantic enrichment of the elements in the learning domain should make possible the automatic creation of workflows by the composition of appropriate services.

3 Paradigm shift

To fully realise the potential of semantics in the pedagogical domain (as described above) it is necessary to make a paradigm shift in the way we deal with semantics [6], this shift happens in two ways:

- From Metadata to Ontologies
- From Information to Knowledge

Ontologies are a more sophisticated way of modelling *metadata*, and *knowledge* is relevant *information* delivered at the right time and context.

3.1 From Metadata to Ontology

Metadata has been widely used to structurally describe learning resources so that they can be better reused. Example standardizations are the Dublin Core [7], which is a general

purpose metadata standard, the IMS Metadata and IEEE LOM [8] (Learning Object Metadata) standard.

While metadata is a starting point to describe the content, recent development in the Semantic Web inspires the use of ontologies for richer semantics. An ontology is “a specification of a conceptualisation” [9]. Ontologies can be seen as an improvement over metadata as they formally define not only keywords (as concepts) but also relationships among them. A simple example shows how an ontology is constructed. This simple ontology defines the concepts of student, teacher and course. The relationships are:

- student *assignedWith* course
- teacher *deliver* course

Apart from the “assignedWith” and “deliver” properties that are associated with their corresponding concepts, each concept would also have its own properties like “name”, “course ID”, etc. Ontologies enable us to make the second shift, from information to knowledge.

3.2 From information to knowledge

Using ontologies enables machines to move from dealing with information to dealing with knowledge (well structured, relevant resources, both content and services, available at the right time and context). Knowledge is sharable and reusable.

When a system has a shared ontology it knows how to handle the semantically enriched resources consistently. For example, when a student wants to search for a course, the course query service knows from the shared ontology what the search criteria are, and these will match with course delivery services even if the two services are developed separately and are deployed at different locations on the Grid (maybe by different software developers and running in different operation systems). The services can understand each other by following the shared ontology.

In the next sections we describe the knowledge life cycle and demonstrate how we have used it to analyse and maintain knowledge in the learning domain.

4 The Knowledge Life Cycle

Knowledge means well structured, relevant resources that are sharable and reusable. To ensure this, resources must be associated with rich semantics that are agreed the members of the domain community. The development and maintenance of ontologies that capture this rich meaning is the subject of Knowledge Engineering. In this section we present the different stages of the *Knowledge Life Cycle*, a model that describes how knowledge is captured, applied and reused. In the following section we will then show how we apply this life cycle to the domain of learning.

Figure 1 shows the four main phases of the Knowledge Life Cycle:

- **Knowledge Acquisition (KA):** The first stage is to acquire the knowledge from the domain experts. This can be done in a variety of ways including

scenario construction and interviews. The objective is to develop a domain vocabulary and a sense of the most important concepts.

- **Knowledge Modelling (KM):** The next stage is for this description to be formalised as an ontology. Classes are defined based on the concepts identified in the KA stage and the possible relationships between those classes are specified.
- **Knowledge Annotation:** Once an ontology has been defined it is tested through application. To do this example resources from the domain are annotated with the ontological metadata. This enables the KM stage to be evaluated and revised.
- **Knowledge Reuse:** Reuse is achieved when new applications reuse the resources (made possible by the shared ontology), for example by incorporating existing learning objects into a new course design.

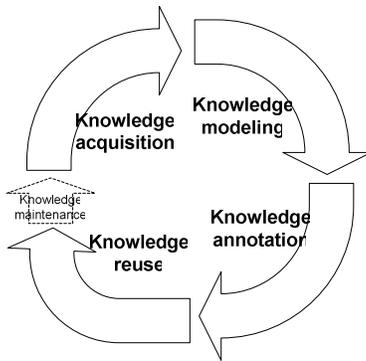


Figure 1: The Knowledge Life Cycle

The Knowledge Life Cycle is intended to iterate over several generations. This means that the ontologies are expected to evolve, and maintenance is necessary. Annotations from earlier generations will need to be updated in order for their reuse to continue. Doing this automatically is still the subject of much research [10] but the formal nature of the life cycle (and its audit trail) ensures that this is at least manually possible.

5 Technical View of Learning Semantics

In this section we will describe how we have used the knowledge life cycle to develop reusable semantics for learning. It is worth repeating that this is not an attempt to create a definitive ontology, but a demonstration of the life cycle within the domain of learning.

Throughout this work we used key mark-up languages, such as XML, RDF and OWL ¹, which are often chosen to represent semantics via ontologies. This formalised expression makes the ontologies machine accessible and interpretable. Figure 2 shows how we interpreted the knowledge life cycle for the learning domain.

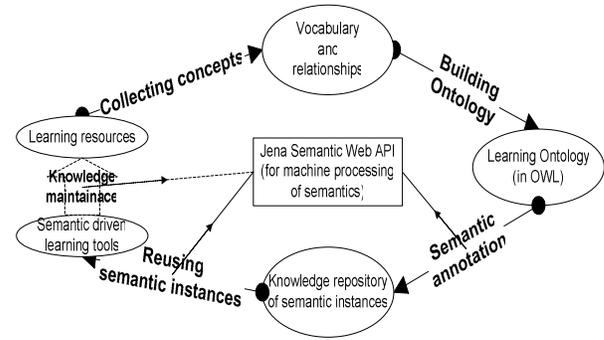


Figure 2: Semantics aspect of the knowledge life cycle

5.1 Building a learning vocabulary

The first part of the Knowledge Life Cycle is the process of Knowledge Acquisition, which has the objective of forming a formal, explicit and shared consensus in the domain. In the learning domain this translates to a process of interviewing learning domain experts and examining teaching and learning materials in order to create a common vocabulary and identify key concepts.

For our own knowledge acquisition the authors examined the output of the OTM 2004 Workshop on Ontologies, Semantics and E-learning [13] (including presentations and papers) and also looked at the online course resources at the University of Southampton (including course notes and syllabuses) and created an initial concept graph. This was then verified with a domain expert in a series of interviews, which resulted in a list of key learning concepts, attributes and relationships. Table 1 shows some examples:

Concepts	Attributes	Relationships
Course, Person, Student, Teaching_Expert, Learning_Event, Poll, FAQ, Institution, University	Teaching_area, design_course, name, description, course_ID, subject	Course – name(1:1); Student – person (is_a); Student – teacher (N:N)

Table 1 Collecting information on the learning domain

5.2 Building Learning Ontologies

The next part of the Knowledge life cycle is Modelling. At this stage the key concepts and terms identified at Acquisition are formalised into an ontology.

Protégé 2000 [11] is an ontology building and knowledge acquisition tool that has been frequently used for knowledge modelling purposes. It allows knowledge engineers to focus on modelling without worrying about the underlying language and syntax. The modelling work can be saved in various formats including RDF and OWL.

¹ XML, RDF and OWL are W3C standards, which can be found in <http://www.w3c.org>

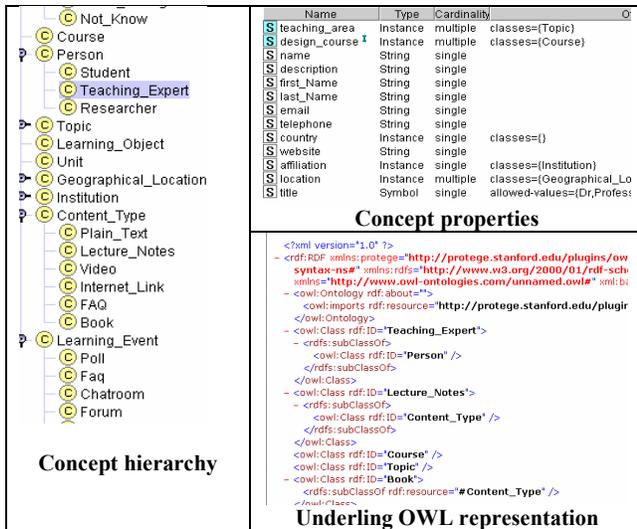


Figure 3 Building domain ontology in Protégé

As can be seen in **Error! Reference source not found.**, we built our initial ontology in Protégé with an OWL plug-in. “Person”, “Topic”, “Learning_Event”, etc. are key concepts under which the taxonomy is further expanded to express hierarchical relationships (parents/children) among concepts. Each concept also has its properties defined to express the subject/predicate relationship (who uses who). The ontological information is saved in OWL format for content enrichment through instance generation.

5.3 Annotating Course Resources

The next step is Knowledge Annotation, this is the process of binding together relevant learning resources with instances from the ontology so that raw content is enriched with more formal meanings pre-defined in the shared ontology. This is also termed *knowledge binding* [12], and depends upon human effort to tag the resources.

Generating the instances involves annotating the raw data source using pre-defined ontologies. For this work we used two methods are used to generate instances. Based on their operational mechanism they are called “Ontology Instantiation” and “Resource Annotation” respectively.

5.3.1 Ontology Instantiation

Protégé can also be used to instantiate an ontology. It may then be treated as a knowledge base or the instances can be saved as independent files.

Figure 4 Generating semantic instances in protégé

shows Protégé being used to create course instances based on relevant information in the original resource (such as its syllabus).

Each instance (in the middle column) represents a course instance. Its properties (“Authorship”, “Prerequisite”) as defined in the ontology) are also filled with object instances, the class of which is constrained by class properties defined

in the ontology. The object instances can be created on the fly or selected from previously generated instances.

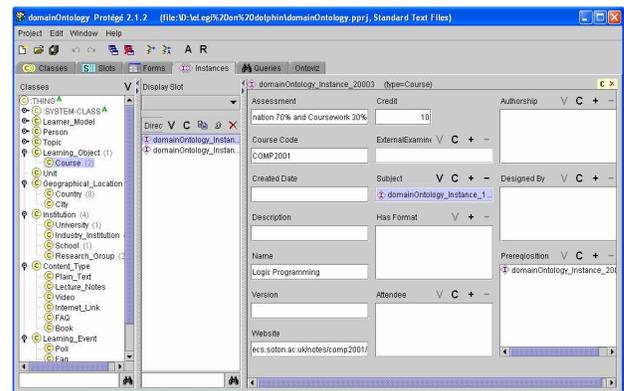


Figure 4 Generating semantic instances in protégé

Instances generated in this way can be exported from protégé (with OWL plug-in) as can be illustrated in **Error! Reference source not found.**, where the instances are represented using RDF as well as OWL enhancement for extra semantics about constraints, for example limiting the cardinality of relationships (in Figure 3 the attributes of the *Teaching_Expert* have a cardinality constraint of either *single* or *multiple*).

5.3.2 Resource Annotation

The task of ontology instantiation is specialised skill that requires knowledge engineers to translate resource information into the ontology, this is often too complicated for resource providers. For the occasions when the resources are generated by these people, in learning this will mainly be teachers and learners an annotation tool would be preferable to allow the end user to do the annotations themselves.

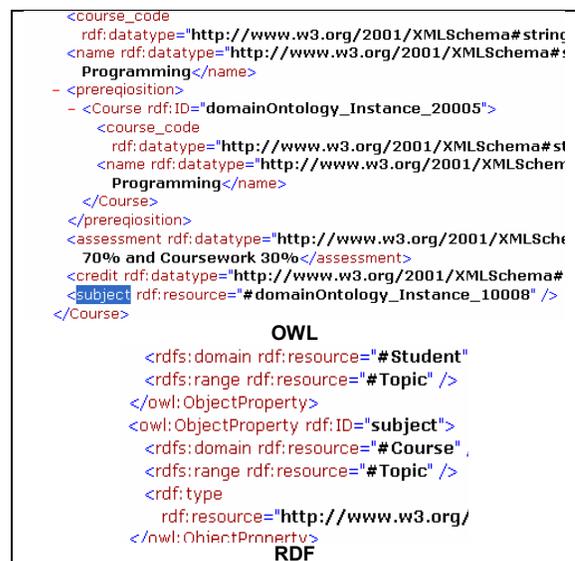


Figure 5 Function semantic instances

5.4 Reusing Course Resources

Once the resources are enriched with semantics, we enter the Knowledge Reuse stage:

1. **Resource discovery:** This is in line with the Semantic Web, which is intended to enrich resources on the Web so that they can be easily identified, located and processed (an example might be to locate a learning object to fulfil a particular course requirement).
2. **Process automation:** As demonstrated by Web Services and the Grid, as services have their interface, parameters and effects semantically described, automation becomes possible (an example might be an assessment service that automatically grades sets of questions).
3. **Service integration:** This is about exploiting semantics to assist the service oriented architecture where simpler services can be combined together to realise more complex customised functionalities (an example might be a live course system automatically assembled and run based on a learning design).

Our intention is to pursue these three reuse objectives, with the first two acting as stepping stones to the third.

6 Conclusion

In this paper we have looked at the semantic aspects of learning from two perspectives: the pedagogical view and the technological view. More sophisticated semantics can enrich learning resources and enable the paradigm shift from information based content delivery to knowledge based, context-aware collaborative learning services. Ontologies can be used as an improvement over existing metadata efforts to bring in the semantics needed for these enriched services and resources.

We have also proposed the use of the Knowledge Life Cycle to manage the key phases in modelling learning semantics. We have described our efforts to follow a life cycle model within the learning domain – namely by performing an acquisition exercise, building a learning ontology and creating semantic instances in Protégé in order to explore automation and reuse in the future.

The paradigm shift from information to knowledge offers serious advantages to the next generation of distributed learning systems. We believe that a Knowledge Life Cycle model is critical to successfully managing learning and teaching semantics and achieving the goals of resource sharing, collaboration and automation.

7 Acknowledgements

Part of this work was funded by The European Commission under the E-Learning Grid Infrastructure project (ELeGI), IST-0002205, Sixth Framework Programme.

References

1. Luis Anido, Judith Rodríguez, Manuel Caeiro, Juan Santos, 'Towards Standards-driven Educational Content Providers' 2nd IEEE International Conference on Advanced Learning Technologies (ICALT'02), Sept 9-12, Kazan, Russia, 2002
2. Ian T. Foster, 'The Anatomy of the Grid: Enabling Scalable Virtual Organizations', 7th International Euro-Par Conference Manchester on Parallel Processing, p.1-4, August 28-31, 2001
3. Judith Rodriguez, Luis Anido, Manuel J. Fernandez, "How can the Web Services Paradigm improve E-learning?" 3rd IEEE International Conference on Advanced Learning Technologies (ICALT'03), July 9-11, Athens, Greece, 2003
4. Toshio Okamoto, Mizue Kayama, "The Learning GRID for Knowledge Building", 4th IEEE International Conference on Advanced Learning Technologies (ICALT'04) August 30 - September 01, Joensuu, Finland, 2004
5. Marco Ronchetti, Paramjeet Saini, "Knowledge management in an e-learning system", 4th IEEE International Conference on Advanced Learning Technologies (ICALT'04) August 30 - September 01, Joensuu, Finland, 2004 pp. 365-369
6. Bhavani Sridharan, Alexei Tretiakov and Kinshuk, "Application of Ontology to Knowledge Management in Web based Learning", 4th IEEE International Conference on Advanced Learning Technologies (ICALT'04) August 30 - September 01, Joensuu, Finland, 2004, pp 663-665
7. The Dublin Core Metadata Initiative, <http://dublincore.org/>
8. IEEE Learning Object Metadata. <http://ltsc.ieee.org/wg12/>
9. T. R. Gruber. "A translation approach to portable ontologies", Knowledge Acquisition, 5(2):199-220, 1993.
10. A. Maedche, B. Motik, L. Stojanovic, R. Studer and R. Volz, "An infrastructure for searching, reusing and evolving distributed ontologies", in 12th International Conference on World Wide Web, Budapest, Hungary, 2003, pgs 439-448
11. The Protégé homepage, <http://protege.stanford.edu/index.html>
12. Tao, F., Cox, S., Chen, L., Shadbolt, N., Xu, F., Puleston, C., Goble, C. and Song, W. 'Towards the Semantic Grid: Enriching Content for Management and Reuse', Delivering e-Science, UK e-Science All-hand Conference 2003, Nott, UK.
13. OTM 2004 Workshop on Ontologies, Semantics and E-learning (28/10/2004), Larnaca, Cyprus, 2004 <http://www.starlab.vub.ac.be/events/OTM04WOSE/>
14. Guus Schreiber, Hans Akkermans, Anjo Anjewierden, Robert de Hoog, Nigel Shadbolt, Walter Van de Velde and Bob Wielinga, "Knowledge Engineering and Management: The CommonKADS Methodology", MIT Press, Dec 1999
15. Demetrios G. Sampson, Miltiadis D. Lytras, Gerd Wagner and Paloma Diaz (Eds): "Ontologies and the Semantic Web for E-learning", Special Issue of IEE IFETS Journal of Educational Technology & Society Special: 2004, Vol. 7, Issue 4