HAL
open science

The Semantic Aspects of e-Learning: Using the Knowledge Life Cycle to Manage Semantics for Grid and Service Oriented Systems

Feng (barry) Tao, Hugh C. Davis, David E. Millard, Arouna Woukeu

To cite this version:

Feng (barry) Tao, Hugh C. Davis, David E. Millard, Arouna Woukeu. The Semantic Aspects of e-Learning: Using the Knowledge Life Cycle to Manage Semantics for Grid and Service Oriented Systems. 1st International ELeGI Conference on Advanced Technology for Enhanced Learning, 2005, Vico Equense (Naples), Italy. pp.8. hal-00190582

HAL Id: hal-00190582

https://telearn.hal.science/hal-00190582

Submitted on 23 Nov 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The Semantic Aspects of e-Learning: Using the Knowledge Life Cycle to Manage Semantics for Grid and Service Oriented Systems

Feng (Barry) Tao, Hugh Davis, David Millard, Arouna Woukeu
School of Electronics and Computer Science
University of Southampton
Southampton, UK
{ft, hcd, dem,aw1}@ecs.soton.ac.uk

Abstract

The aim of bringing semantics to learning content and services is to enable large scale collaboration of e-learning activities over the Grid infrastructure. As machines and software applications will be ubiquitously involved in enabling and facilitating this collaboration, it requires a common understanding of the domain, in particular at the conceptual level, so that both computers and human participants of the e-learning activities understand and are able to communicate among themselves though a common conceptualisation of the learning domain. In this paper we propose that Ontology is required in this context which is shared as a language to enrich resources with semantics. In line with the pedagogical view of the semantics obtained from experts in the learning domain, we describe a semantics driven knowledge life cycle to address the key phases in managing semantics in e-learning. Service oriented architecture is adopted for the vision of Grid infrastructure, reusable components and easy integration at a later phase.

1. Introduction

Learning related content and services are two main resources that are manipulated and interacted with in the e-learning activities. The concept of semantics is brought in for resource enrichment so that they can be better understood and processed by third-parties. In this aspect, we describe a technical view and a pedagogical view of the semantics. From the technical points of view, Ontology is a specification of a conceptualization. As an improvement over metadata, it is the backbone to better realize the Semantic Web/Grid vision – the idea of having resources on the Grid semantically enriched and linked for more effective discovery, automation, integration, and reuse across various applications.

To start, we give a simple example to illustrate the potential benefit of bringing semantics into the learning domain. The ontology defines the concept 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 also has its own properties like "name", "course ID", etc. Real people (students and lectures) and courses are the raw resources that need to be semantically enriched with the ontology. Ontology driven templates can be automatically generated so that any student/lectures can fill them out and generate corresponding semantic instances that represent a semantic enriched instance of the raw resource of themselves. As the system has a shared ontology, it knows how to process these requests and handle the semantically enriched resources. For example, when a teacher wants to register a new course, the course generation service knows from the ontology what properties of the course it should ask the teacher to fill-out in order to generate a semantic instance of the "course concept". Furthermore we envisage that by adding "sequence" concepts the system could automatically generate the "workflow" that will result from the introduction of this course. On the other side, when a student wants to search for a course, the course query service knows from the shared ontology what the search criteria are. By this means, the two services deployed at different locations on the Grid (maybe by different software developers and running in different operation systems) can understand each other by following the shared ontology. Furthermore, the Declarative Logic (DL) based inference allows students to classify their courses and select potentially beneficial ones automatically based on the inferred facts.

The Grid is now a well established infrastructure that underpins the above purposes very well. Firstly it allows resources (content and services) to be located in a geographically distributed environment yet the semantic aspects can still guarantee the seamless integration and process automation. Secondly, it provides a potentially large pool of CPU and memory to handle the computationally intensive DL reasoning for inferring new facts automatically from existing facts stated in the semantics.

The vision is that a portal can be architected to embrace a distributed, service orientated architecture, where different parties can log in, discover, share, integrate and collaborate on learning resources. These activities are driven by a set of predefined ontologies and their upper level semantics bond with the resources.

1.1 Purpose and scope

This section aims to help us agree a common understanding of what it means to add semantics to the grid, and to the learning grid in particular.

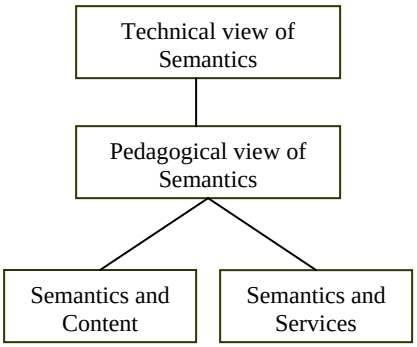


Figure 1 Research aspects of semantics in e-learning

As illustrated in
Figure **1**, the semantics can be brought in to describe resources, which can be categorised into two major types – the content and the services in Grid-based e-learning.

On the other hand, the semantics can also be examined from a pedagogical view and from a technical view. The technical view on Semantics focuses on generic specifications and technologies on Ontology, Semantic Web, and Semantic Grid, etc. The pedagogical view on Semantics places more emphasis on concrete impacts that semantics bring to distance education or e-learning activities.

2. Pedagogical view of semantics

The purpose of this section is to examine the affordances of semantics from a pedagogical point of view, with the intention of motivating the semantic layer of the ELeGI architecture. We are not interested here in how the semantics may be implemented but in the uses to which semantic mark-up can be put in order to benefit and improve learning and the formal and informal processes surrounding learning. The implementation of metadata and semantics for learning is covered in [2], particularly sections 5,6,7 and 8, and in [3].

2.1 The constituent elements of learning

From a learner's point of view there are a number of interactions which may take place:

- learner ←→ *teacher*
- *learner* ←→ *expert*
- *learner* ←→ *learner*
- *learner* ←→ *resource* (e.g. a book or a web page)
- *learner* ←→ *service* (e.g. a search engine, an on-line experiment, a course schedule)

It is not our intention here to get into profound arguments about the meanings of these terms. We use the term teacher to mean a person who has a formal role in guiding student's learning, an expert as a person who may help a learner but without the formal role, and the *learner* ←→ *learner* interaction acknowledges that learning will often happen in discussion with other learners; the distinction between these three roles is not fixed. Similarly we characterize the difference between a resource (something which is essentially static and should not change between accesses) and a service (a dynamic process producing results varying according to context). Again, there is no hard line between these elements; we are using the terms in order to aid communication and understanding in the later discussion rather than as a formal taxonomy of learning.

We should also consider the role of those whose job is to support learning. Often this role is taken by teachers, but may also include librarians, learning technologists, web authors, and technicians. ELearning practitioners often comment that they believe they spend as much time organising the materials, setting access permissions, dividing students into working groups,

setting access times etc. as they spend on teaching and the production of teaching materials.

We suppose that in the future systems are going to deliver greater intelligence in order to improve learning and increase the efficiency of the management of learning. In order to achieve these goals, clearly we it will be necessary for systems to have an understanding of the learners, the teachers, the domain experts, the resources and the services.

2.2 What are the services?

The E-Learning Framework [4] provides a useful mapping of the functionality of E-Learning Systems onto services. It distinguishes between three layers of services:

Common Services: These provide the base level system functionality on which other services rely and which are not really specific to the learning domain. Examples are *Logging, Alert, Format Conversion, Forum, Chat, Authentication, Authorization, Metadata Management, Content Management* and *Federated Search*. These same set of services will be required by any middleware which attempts to provide an environment in which people can work co-operatively in virtual communities.

Learning domain services: These are there services which are specific to the formal learning domain. They are split broadly into four sub themes,

- Course Creation (e.g. *Curriculum, Course Management, Activity Author*)
- Course Delivery (e.g. *Activity Management, Learning Flow, Sequencing*)
- Assessment (e.g. *Marking, Grading*)
- Record Keeping (e.g. *Reporting, ePortfolio, Personal Development, Quality Assurance*)

User Agents: These are the interfaces through which the users will interact with the other services. Examples include *VLEs, LMSs, Portals, Library Systems, Assessment Tools, Assignment Marking Tools, Authoring Tools, and Timetabling Tools* etc.

It would be fair to say that the functionality of most of these services already exists in the current monolithic Managed Learning Environments such as Blackboard and WebCT. There are many advantages to adopting a service oriented architecture to supply such functionality (support for pedagogic diversity, composable, flexible functionality, easier collaboration between institutions etc. [12]) but the issue we are concerned with here is how adding semantics to the constituent elements of learning (*learners, teachers, domain experts, resources* and *services*) can improve the learning environment and the learning itself.

The most obvious improvements of having well defined metadata are the opportunities for reuse of learning materials, the ease of interoperability and the general improvements in the ease of accurately searching for information. However, in the following sections we will attempt to move beyond the obvious to describe examples where intelligent reasoning applied to the semantics adds novel value to the pedagogic endeavor.

2.3 How semantic enrichment can improve learning

In this subsection we briefly describe a number of the ways in which intelligent reasoning about semantics can improve the opportunities for a student to learn. Most, but not all, of these scenarios are specifically concerned with formal learning; they address the whole educational cycle from studying, researching, assessment and feedback.

- **Connecting communities**: Services can put people in contact with other people who are experts or learners with similar interests. This is a particularly important tool in informal learning scenarios.
- **Personalized content**: Intelligent tutoring systems have for some time being delivering content that was personalized for the user, based on an understanding of their goals and previous knowledge.
- **Personalized sequencing**: Adaptive Hypertext Systems attempt to provide pathways through materials by matching domain ontologies with dynamically evolving user models.
- **Adaptive Assessment**: Systems may choose questions for the learner that attempt to test at just below the boundary of their knowledge and understanding, thus improving the efficiency of assessment and providing feedback that provides detail in the area that the student can most use it. Such assessment can be in the form of “self test” questions (formative assessment) or more formal assessments used for the purpose of grading (summative assessment).
- **Feedback Agents:** Intelligent agents that observe student behaviour (e.g. assessment results, interactions with a virtual experiment etc.) can

attempt to provide feedback and links to suitable material to assist the learner.

- **Recommender Agents**: Intelligent Agents that observe the information searching and studying patterns of users may recommend alternative materials. In a formal learning setting, the agent could query the syllabus and the course timetable to recommend a plan of study.
- **Annotation Tools**: Allowing users to annotate information is an excellent source of semantic information providing useful information for others and allowing both readers and other services the opportunity to process the information in alternative ways (e.g. a search engine might be able to find all commentaries by a critic of a given author's work).
- **Search Engines**: When information has been semantically enriched then search engines can be much more powerful, having a much better ability to isolate relevant parts of documents and the ability to identify a much greater range of media types (videos, on-line experiments, simulations, data feeds, etc). Where services are semantically enriched, then search engines can choose suitable services to manage the query.
- **Analytic Tools**: The e-Science community is leading the way in production of tools that harvest, store and analyse data from a range of sources.

2.4 How Semantic Enrichment can improve the Management of learning

- **Production of materials**: Production of teaching materials is a notoriously time consuming task, and the ability to locate and to re-use existing materials is a primary motivation for providing metadata for learning resources. The next stage is to provide services to assist in the location of suitable materials from heterogeneous sources.
- **Student Management**: An understanding of the roles of the actors (teachers, students, experts, 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 some new materials, the closing date of some assessment, the exact time of a synchronous group chat session etc. These events can be made to happen automatically when a course is described in some language such as IMS Learning Design.
- **Record Keeping**: Record keeping and quality assurance can be the bane of a teacher's life, requiring them to spend much time ensuring that all the results are kept in the correct places such as institutional enterprise systems, student portfolios as well as made available for QA purposes by whatever external authorities might be involved. All of this work is an obvious target for automation by services that understand the goals.
- **Quality Assurance**: Quality assurance often involves the maintenance of sample work and feedback/reflections, as well as ensuring that new programmes, courses and assessments have been through appropriate validation. Again, this is a task that could be assisted by intelligent services which could guide such tasks through the set of other services involved.

Much of what has been described in this section is in effect suggesting the use of some kind of agent to orchestrate 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, then might call an email service to inform the students of the need to update their personal development plans accordingly. This sort of orchestration of services looks very much like "workflow". It is our belief that appropriate semantic enrichment of the elements in the learning domain should make possible the automatic creation of such workflows by the composition of appropriate services.

3. Knowledge life cycle of semantics

In order to achieve these visions, we proposed a Semantic Web based knowledge management approach. We illustrate in Figure 2 the key steps in the knowledge life cycle as well as tools/APIs that we intend to use in each step.

Knowledge acquisition is carried out to acquire key information in the learning domain, the result of which is then fed to ontology modeling where we use the protégé ontology tool with an OWL plug-in to build a formal ontology. Ontology forms the conceptual structure of the knowledge base and the semantic annotation populates the knowledge base by generating instances. This is done by binding content with the relevant part of the ontology. The ontology and the semantics (instances) are stored in OWL format for easy machine processing at the programming level.

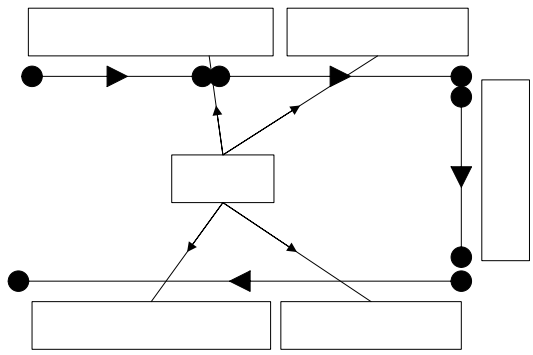


Figure 2 Semantics aspect of the knowledge life cycle

these instances to generate knowledgeable decisions. Two types of reuse are envisaged:

1. Ontology driven query of the learning resources: dynamic forms can be automatically generated according to the pre-defined ontology so that end users know what the query criteria are.

2. Semantic matching based knowledge advisor on course recommendation: recommend and rank courses according to course ontology and learners' profile data. E.g., semantics on course pre-requisitions, learner's level, learning curve and subject, etc.

Note that the ontology is the share consensus and has been stored in the OWL format, Third-party programs must be able to access it and process the instances in the knowledge base for different knowledge reuse purposes. In order to allow the knowledge to be used potentially outside the awareness of its providers, a generic OWL API must be used. We use the Jena semantic toolkit [6] to process the OWL format ontologies and semantic information.

4. Technical view of semantics

Content enrichment is an important step in the semantics life cycle from knowledge acquisition to reuse. It binds together relevant parts between content and ontology so that raw content is enriched with more formal "meanings" pre-defined in a shared ontology.

Enriching content is also termed as knowledge binding [11], which depends upon human effort to tag thus making the knowledge accessible.

Key mark-up languages, such as XML, RDF and OWL [1]are often chosen to represent the semantics via ontologies and make them machine accessible and interpretable.

This section will be informed by sections on Learning Object standards from ELeGI deliverable D6 (Projects, technologies and Standards evolution monitoring report) and the SEES where appropriate.

We address here how e-learning activities benefit from semantically enriched content.

4.1 Ontology Engineering

Prior to the enriching phase, the ontology must be made available as a formal, explicit and shared consensus in the domain. We intend to build the ontology by working together with domain experts in e-learning and utilize existing key specifications of content representation as described in the previous section.

Ontology is a specification of conceptualization [13]. It explicitly defines the domain concepts and their relationships. It is similar to a dictionary or glossary, but with richer structure, relationship and axioms that describe a domain of interest more precisely.

As can be seen in Figure 4.1, we build an 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 (parent/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.

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

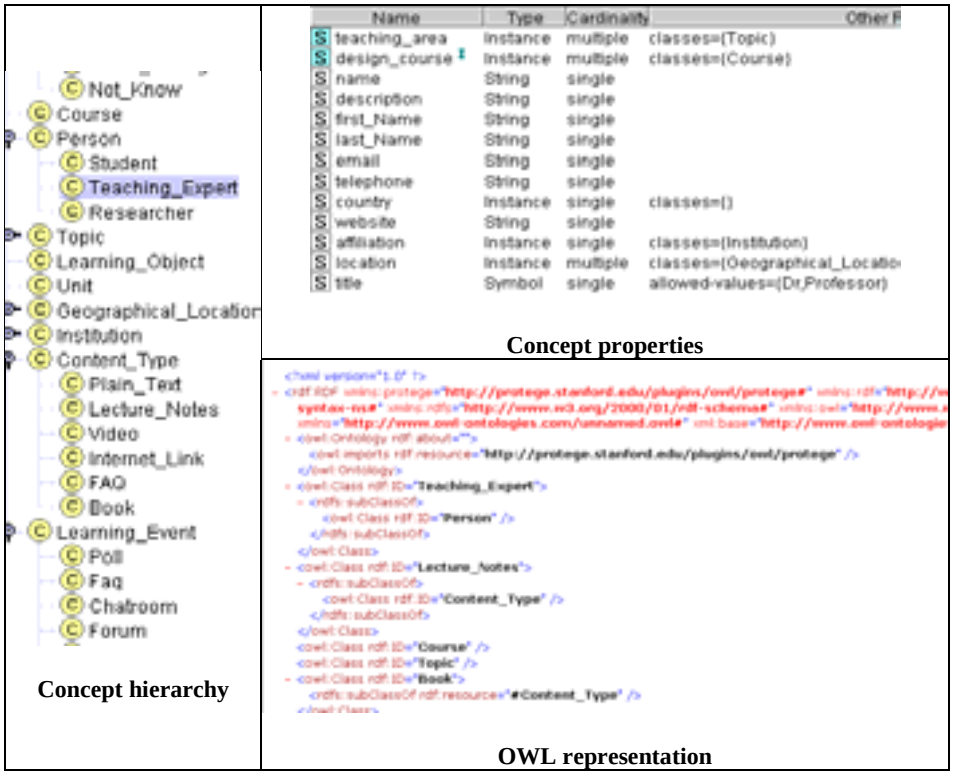


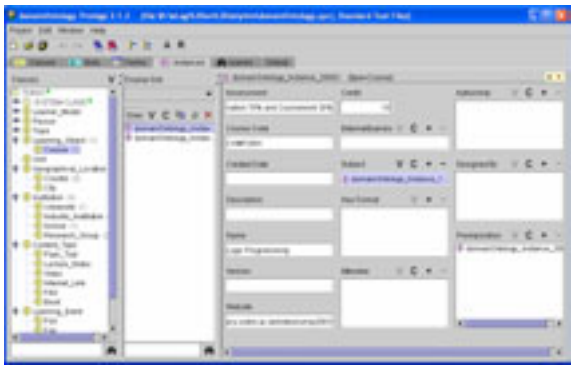
Figure 4.1 Building Ontologies

4.2 Enriching resources

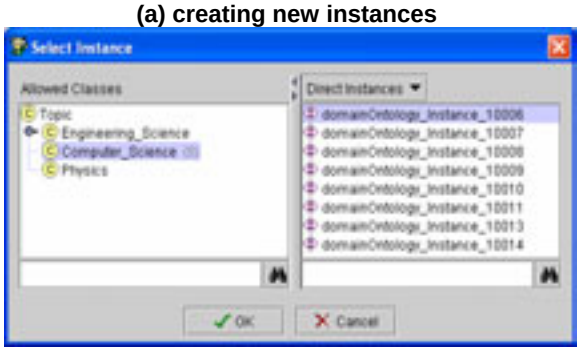
While ontology is important in specifying the conceptual structure and a constrained vocabulary set, instances are treated as the concrete content in a semantic knowledge base. Generating the instances involves annotating the raw data source using pre-defined ontologies. In this paper, two methods are used to generate instances. Based on their operational mechanism, they are called "Ontology Instantiation" and "Resource Annotation" respectively.

1) Ontology instantiation

Protégé 2000 [9] 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 (given OWL plug-in has been used).



(a) creating new instances



(b) selecting subject instances

Figure 4.2 Generating semantic instances in protégé

As illustrated in Figure 4.2-a, to create semantic instances of courses, relevant information in the course content is used to instantiate its corresponding ontology classes, such as "*Assessment*", "*Course code*" and "*Subject*", as defined in the ontology in Figure 4.1. Each instance (in the middle column) in Figure 4.2–a represents a course instance. Its properties ("*Authorship*", "*Prerequisition*" 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.

In order to reach the maximum capacity of expressiveness, many languages have been designed to express the ontology and semantic information. Among them, the most recent is the Web Ontology Language (OWL), which has evolved from RDF and DAML+OIL to provide more expressive power. RDF stands for Resource Description Framework, a graph model (triple statement) which is designed for describing and searching resources on the web. The DAML+OIL is a schema language that adds constraints on properties to assist machine reasoning. For example when "daml:TransitiveProperty" is added as a constraint on the property "P1:older_than" of a RDF model, if we have A1:P1:A2 and A2:P1:A3, then A1:P1:A3 can be inferred. This is very useful for reasoning and inferring new knowledge that has not been directly stated. DAML+OIL also uses subProperty to describe relationships at different granularities. OWL is close to DAML+OIL with additional constraints such as "sameAs", "cardinality", etc. for more expressive power.

Instances generated in this way can be exported from protégé (with OWL plug-in) as can be illustrated in Figure 4.3, where the instances are represented using RDF as well as OWL enhancement for extra semantics. The RDF can be also interpreted as N-Triples for efficient machine processing.

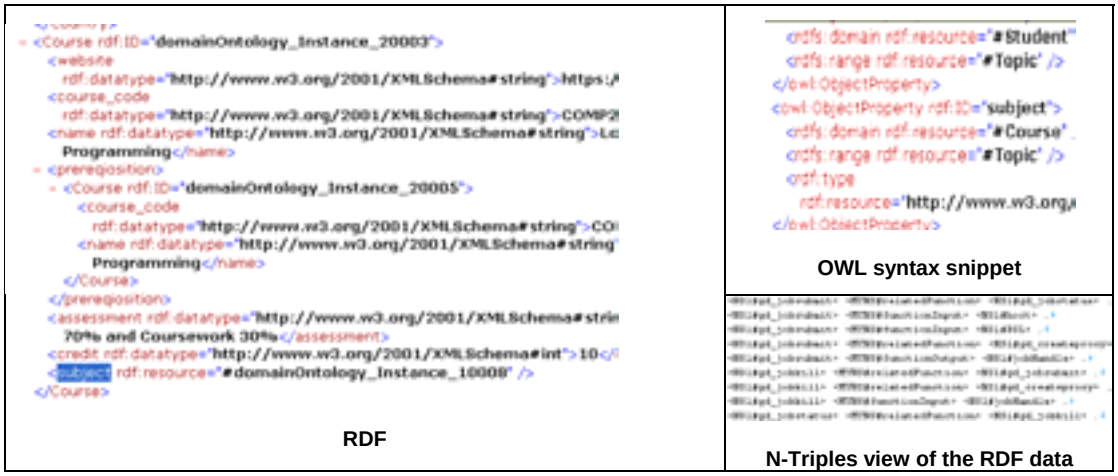


Figure 4.3 Function semantic instances

2) Resource annotation

While in Protégé, knowledge engineers acquire information about resource to instantiate an ontology, this is often too complicated for resource providers. On the other hand, there are many occasions when the resources are actually generated by the learners. In this case, a annotaion tool will be preferable to allow the end user to semantically annotate thier own resources for reuse.

Learning services are another important resource that need to be annotated so that we can be easily located for service orchestration. OWL-S (evolved from DAML-S [14]) is a standard language initially designed to semantically enrich services. Further development can be found in [13].

4.3 Advanced semantics reuse

Once the resources are enriched with semantics, more advanced reusing can be carried out in the form of discovery, automation and integration. At the moment, these are only our visions. More work need to be done to demonstrate these visions

1. Resource discovery: This is in line with the Semantic Web and Semantic Grid, which intend to

enrich data/services on the Web/Grid so that they can be easily identified and located for reuse purpose.

2. Process automation: As services /content have their interface/structure semantically described, automation is possible by using agent technologies.

3. Service integration: This is about exploiting semantics to assist the service oriented architecture where simpler services can be orchestrated together to realize more complex customized functionalities.

5. Conclusion

The semantic aspects are viewed from two perspectives: the technological view and the pedagogical view. We argue that the semantics can be brought in to enrich learning resources – both content and services in the learning Grid. Key technologies such as ontologies, the Semantic Web, the Semantic Grid and web services are intended to contribute to a new paradigm where the learning environment becomes service oriented, context aware and more user friendly.

Experience from Semantic Grid research in Geodise [5] illustrated that although ontology building is the kick-off activity for semantics based knowledge management, ontologies manipulation and maintenance activities happen all the way toward the end of the knowledge management life cycle.

In particular, the design of the semantics should be motivated by smart knowledge reuse scenarios that intend to facilitate e-learning related resource sharing and collaboration.

References

1. [Duval2004] Erik Duval, Towards a More Even Distribution of the Future: Insights from ARIADNE: Keynote at Merlot Conference, 2004 [http://conference.merlot.org/conference/2004/presentations/20040804_Merlot_CostaMesa.ppt]
2. [ELeGI deliverable D6] - 1st Projects, technologies and Standards evolution monitoring report. ELeGI Internal Report: 2004
3. [ELeGI deliverable D10] - Preliminary version of Didactical and Knowledge representation models for VSE. ELeGI Internal Report: 2005
4. [ELF] The E-Learning Framework: [http://www.elframework.org/] Accessed 06/02/05
5. [Geodise] Grid Enabled Optimisation and Design Search for Engineering, e-Science polit project, http://www.geodise.org
6. [HPJena] HP Labs Semantic Web Research, http://www.hpl.hp.com/semweb/
7. [McCalla, 2004] McCalla, G. The Ecological Approach to the Design of E-Learning Environments: Purpose-based Capture and Use of Information About Learners. Journal of Interactive Media in Education, 2004 (7). [www-jime.open.ac.uk/2004/7]
8. [Nejdl & de Bra, 2004] Nejdl, Wolfgang; De Bra, Paul (Eds.) Adaptive Hypermedia and Adaptive Web-Based Systems: Third International Conference, AH 2004, Eindhoven, The Netherlands, August 23-26, 2004, Proceedings Series: Lecture Notes in Computer Science, Vol. 3137 2004, XIV, [http://www.springeronline.com/sgw/cda/frontpage/0,0,3-0-22-33913464-0,0.html]
9. [Protégé] The Protégé homepage, http://protege.stanford.edu/index.html
10. [Sampson et al., 2004] 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; [http://www.ifets.info/]
11. [Tao2003] Tao, F., Cox, S., Chen, L., Shadbolt, N., Xu, F., Puleston, C., Goble, C. and Song, W. (2003) Towards the Semantic Grid: Enriching Content for Management and Reuse. In Proceedings of Delivering e-Science, UK e-Science All-hand Conference 2003, Nottingham, UK.
12. [Wilson et al. 2004] Scott Wilson, Kerry Blinco & Daniel Rehak, Service-Oriented Frameworks: Modelling the infrastructure for the next generation of e- Learning Systems. July 2004. [http://www.jisc.ac.uk/uploaded_documents/AltilabServiceOrientedFrameworks.pdf]
13. T. R. Gruber. "A translation approach to portable ontologies", Knowledge Acquisition, 5(2):199-220, 1993. http://ksl-web.stanford.edu/KSL_Abstracts/KSL-92-71.html
14. OWL-S, http://www.daml.org/services/owl-s/1.0/
15. A.. Ankolekar, M. Burstein, J.R. Hobbs, O. Lassila, D. Martin, D. McDermott, S. A. McIlraith, S. Narayanan, M. Paolucci, T. Payne, K. Sycara. *DAML-S: Web Service Description for the Semantic Web.* In the Proceedings of the 1st International Semantic Web Conference, Sardinia, Italia, June, 2002.