



HAL
open science

Ontologies and Knowledge Aggregation in the Active Semantic Learning System

Ioan Szilagyi, Radu Balog-Crisan, Ana Roxin, Ioan Roxin

► **To cite this version:**

Ioan Szilagyi, Radu Balog-Crisan, Ana Roxin, Ioan Roxin. Ontologies and Knowledge Aggregation in the Active Semantic Learning System. 11th IEEE International Conference on Advanced Learning Technologies (ICALT), Jul 2011, Athens, United States. hal-00701062

HAL Id: hal-00701062

<https://hal.science/hal-00701062v1>

Submitted on 24 May 2012

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.

Ontologies and Knowledge Aggregation in the Active Semantic Learning System

Ioan SZILAGYI; Radu BALOG-CRIȘAN; Ana ROXIN; Ioan ROXIN
Laboratory ELLIADD, University of Franche-Comté
Montbéliard, France
{firstname.lastname}@univ-fcomte.fr

Abstract — The construction of semantic-based learning systems depends on the development of ontologies and the capacity to integrate and exploit knowledge using semantic technologies, notably RDF and ontologies. In this paper we present some of the concepts and ontologies defined in the context of the Active Semantic Learning System (Active SLS) that are used to describe resources and the semantic relations between the concepts defined in the different ontologies. The purpose is to obtain a learning system that is capable of aggregating knowledge from different sources from the web and of exploiting that knowledge for the benefit of the learner.

Keywords: Active SLS; semantic learning; ontologies; knowledge; RDF; OWL; LOM; User Profile

I. INTRODUCTION

Using semantic technologies in web-based learning systems has been in the interest of researchers for a while now. In this paper we present an approach allowing learning systems to benefit from new web technologies in order to improve the quality, effectiveness and efficiency of the learning process. Constructed using semantic technologies, the Active Semantic Learning System (Active SLS) takes advantage of Semantic Web and Social Semantic techniques to deliver consistent and pertinent learning resources to the learner, while encouraging learners' active participation through the learning process.

The Active SLS aggregates sources of knowledge in order to construct a learning specific knowledge base that can be exploitable for the benefit of learner. That knowledge base together with a learning management should help the learner in his learning process, providing him with consistent and adapted learning resources. The knowledge is contained in different repositories over the web and the system architecture allows the integration of user-generated content as learning material as supplement for learning tasks.

At the system level, knowledge is expressed using semantic web languages like RDF (Resource Description Framework) and OWL (Web Ontology Language).

Regarding the type of information, different modules compose the Active SLS. Each module can be used separately from the system and manages information specific to a domain. The management of the learning process is assured through ontologies and rules defined at application level. The application is in development and not all functionalities have been tested.

Other approaches already exist, [1][2][3], and they all point to the fact that learning process will be improved from the integration of social semantic web in the learning systems.

II. LAYER ARCHITECTURE OF THE ACTIVE SEMANTIC LEARNING SYSTEM

The system is organized into three different levels according to the level description of data as presented in Figure 1. In order to be able to identify and to process the connections between resources, these have to be described using semantic technologies. The transformation from XML, or relational databases, to RDF is assured by applications like GRDDL and D2RQ. If the data already exists in RDF format, the system can take advantage of it. If the resources do not contain in their original repository enough metadata in order to be exploitable in the learning system, professors or learners can add these later and manually.

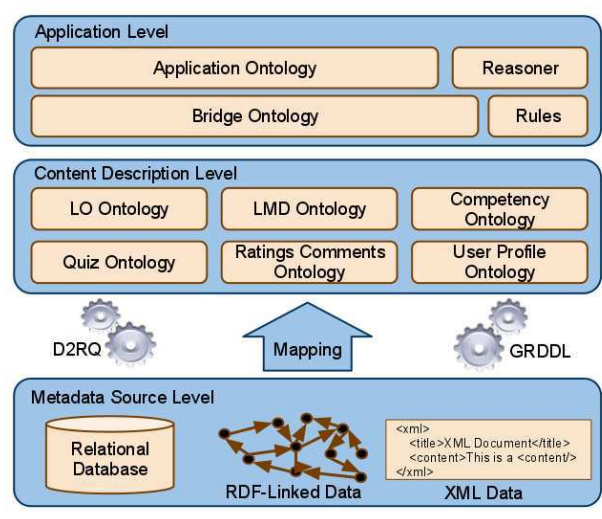


Figure 1. Layer Architecture of the Active Semantic Learning System.

The data flows from the Metadata-Source Level to the Description Level and finally to the Application Level where the system can benefit from it.

A. Metadata-Source Layer

The Metadata-Source Layer actually constitutes the sources of metadata. The sources exist on the web and, in order to be exploitable, they must be accessed. The data and meta-data are present in different forms and formats. The meta-data is used to integrate the resources in the Active SLS, to make it exploitable by the system and it is mapped to an existing model according to the resource type. Universities are already making their content openly

available on the web through different services like: MIT (OpenCourseWare), Stanford (YouTube).

Some of the information regarding the learning courses from OpenCourseWare, which offer a Metadata API [10] for searching in their repositories, is:

- CourseURL: The URL of the course's home page.
- CourseLanguage: The language of the course.
- Title: The course title.
- CourseSection: The department or faculty of the course as given by the institution. If a course is cross-listed in more than one department, this lists only the actual teaching department specified by the institution.
- Description: The text of the description of the course.
- DownloadPageLink: The URL of the page that contains the download link of the course.

YouTube, which offers an API [16] to integrate their content, includes the following metadata regarding video content: Id, Published, Updated, Category, Title, Content, Link, Author, Comments, Location, Media (category, content, description, keywords, player, thumbnail, title, duration) Ratings, Statistics. The YouTube API provides functionalities in order to launch a video at a desired time. The learner is not obliged to watch an entire video.

Metadata offered in these cases by these APIs is not sufficient in order for the system to fully exploit the learning object in different usage scenario. Manually intervention is needed from the part of the professors and students.

Other resources can be integrated in the system. Resources that use semantic technologies (e.g. DBpedia) can be exploited directly at the Description Layer.

B. Description and Application Layer

One of the advantages is that knowledge held in different repositories is aggregated and used by the system.

The Description Layer contains metadata modeled and mapped according to ontologies defined in the context of the module. Professors and learners can enrich metadata around different resources in order to better express the value and the content of resources. At this level, knowledge is modeled in order to be compatible and aggregated with other sources.

Knowledge aggregation is more than information retrieval. The system must integrate the information along with other sources and form a local data model that has relations between the data and information it includes.

The Application Layer contains metadata defined at the Description Layer, but merged and viewed as a single graph. Different models are aggregated and form a big graph with the correspondence defined in the Bridge Ontology. The application can execute queries over the graph and obtain results from all implied modules and ontologies.

III. DEFINED ONTOLOGIES

Several ontologies are defined in order for the application to manage and aggregate knowledge. It can be viewed like a single large ontology, but for management purposes we

divided into several ontologies around main actors and concepts in the learning process. The concepts modeled are structured into the following ontologies:

- LMD (Bachelor-Master-Doctorate) models administrative concepts and learning process used in universities
- User Profile Ontology is the description of the learner in the learning context, with his learning objectives, skills and achievements.
- LO (Learning Object) Ontology is used to describe the units of knowledge and the learning objects used in learning process.
- Competency Ontology defines the achievements and learning objectives of the learner, but also the scale on which one can measure the achievements of the learner.
- Quiz Ontology provides context for the evaluation and verification of the learner regarding knowledge achievements.
- Appreciation Ontology is used to appreciate different resources from the system: learning resources, comments, posts.

In all defined ontologies there exist terms and elements imported from different ontologies recognized in their domain of application: DCTERMS [12], FOAF [13], SIOC [14], SKOS [15]. In the following section, we present the different ontologies, their purpose together with some concepts they define.

A. LMD Ontology

Part of the LMD Ontology is illustrated by Figure 2. . This ontology provides context and vocabulary for describing administrative context of the learning process. The main classes are: Student, Professor, Specialization, Course, LOPath.

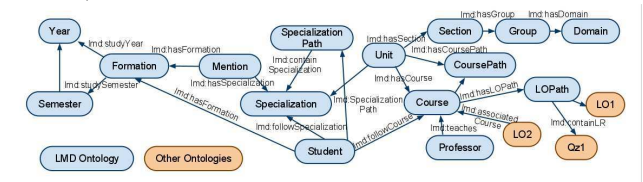


Figure 2. Part of LMD Ontology.

The student has a specialization and he is in relation with professor through the courses he follows. He can access the learning resources by the association between course and learning resource through the LOPath, or just browsing the course's associated learning resources. The learning-object path (LOPath class) is attached to a course (Course class) and provides information regarding the list and order of learning resources covered by the course.

B. User Profile Ontology

The main actor in the learning process is the learner; he is the beneficiary of the learning process. The User Profile Ontology, as illustrated by Figure 3. , models the learning aspects of the learner, and particularly his learning style.

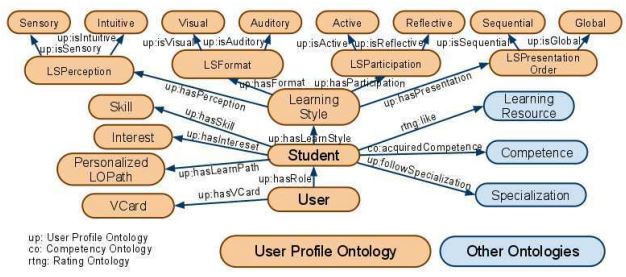


Figure 3. Part of User Profile Ontology

The learner is modeled with the purpose of offering him better learning experiences and content adapted to his needs.

Main classes are User, LearningStyle, PersonalizedLOPath. In our approach, we have chosen to model the learning style as proposed by Felder-Silverman [1]. Interest and Skill classes are defined for better associate the learner objectives and preferences with the learning context and provided content.

C. LO Ontology

The Learning Object Ontology describes the learning objects. The learning objects are units transmitting knowledge, self-contained and reusable. The main classes and properties are those declared by IEEE LOM. The LO Ontology is based on the IEEE LOM Description Metadata. The ontology uses the DublinCore Metadata according to specifications from [8].

The ontology is extended with properties that offer the possibility to relate a learning object to different learning styles.

D. Competency Ontology

The final purpose for the student is to achieve the skills for which he enrolled in the education process. Several researches have been done regarding the modeling and representation of skills. Interest has been showed from the human resource management area [22]. In the learning context, efforts have been made for integrating the skill description into IEEE LOM [6]. In our approach, we exploit the terms and vocabulary defined by these propositions, but we build a separate ontology for better administration.

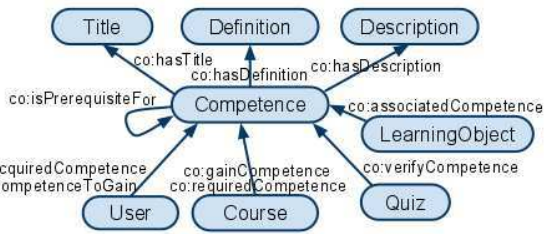


Figure 4. Part of Competency Ontology.

The classes defined in the Competency Ontology, as showed by Figure 4. , are in accordance with the data elements described by the Reusable Competency Definitions [7].

The competencies are in relation with the specializations, units, and courses concepts from the LMD Ontology. At the same time, a learning object can be associated to a skill. The validation of a skill by a learner can be made as a result of validating different quizzes. There is also a hierarchy of skills, underlying the idea that the access to a given skill can be restricted for a learner that doesn't have a prerequisite list of skills.

E. Quiz Ontology

A learner can validate a skill when passing a declared level on a quiz or a group of quizzes. The Quiz Ontology, as illustrated by Figure 5. , offers the context for verifying the skills' achievement.

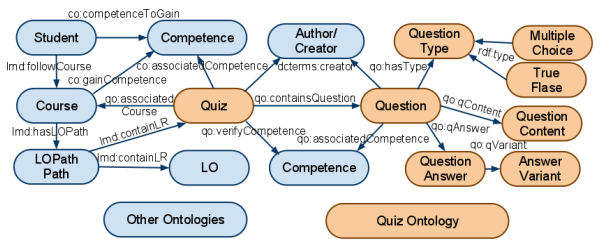


Figure 5. Part of Quiz Ontology.

The central concepts of the ontology are: Quiz and Question. The ontology is described in [5].

F. Appreciation Ontology

The Appreciation Ontology, as showed in Figure 6. provides ways, context and vocabulary for the user to express preference and rate the content. Such noting can be done in several forms: ratings, comments and tags are just a few ways. Several implementations exist for the Tag Ontology and the Ratings Ontology. We preferred to group these ontologies into the Appreciation Ontology.

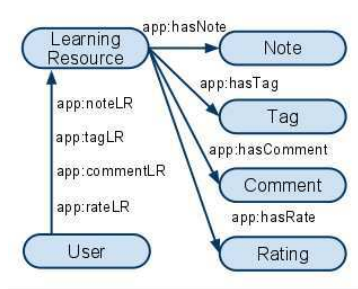


Figure 6. Part of Appreciation Ontology

Among the different initiatives that exist for annotating content, one may cite The Tag Ontology [18] and OATS [9].

Appreciations (notes, tags, comments, ratings) can draw attention and influence the value of a resource, therefore making it more visible for the learners. The Appreciation Ontology can also express the activity of the learner in the system. Different usage scenarios can be envisaged, but for now the Appreciation Ontology is only used to provide another opinion for a learning content.

IV. LINKAGE OF ONTOLOGIES AND KNOWLEDGE AGGREGATION

The Bridge Ontology is defined at the application level. It links the declared ontologies in order to obtain a complete graph, as illustrated by Figure 7. , allowing to infer and to make queries over it.

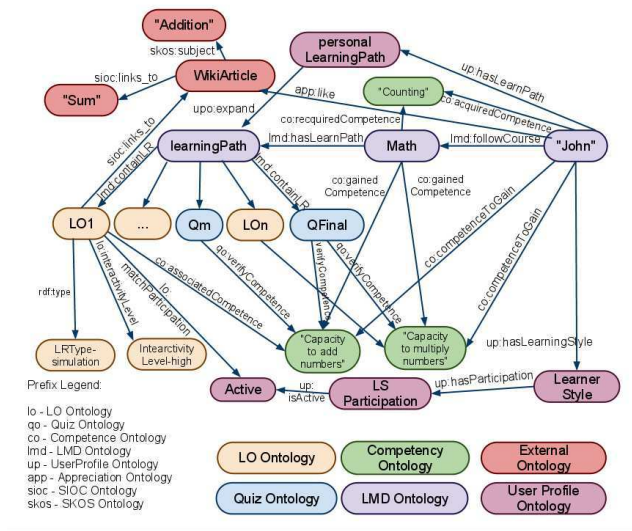


Figure 7. Part of RDF Dataset

Metadata about learning objects can be added to the internal repository by professors or students. Given the poor presence of learning resources' descriptions over the web, in most cases these cannot be inferred or automatically deduced, therefore manual intervention is needed.

A. Application and Bridge Ontology

The Application Ontology is an abstract model of the data that has to be handled at the application level. It contains all vocabulary in order to launch queries over the combined RDF dataset. The Bridge Ontology plays a key role. It merges and maps the ontologies defined at the module level to the Application Ontology. It offers the possibility to integrate and use other externally defined ontologies.

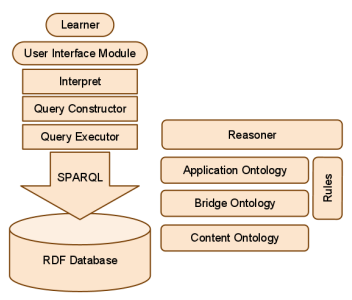


Figure 8. Semantic Kernel – Architecture

These ontologies are declared at the Semantic Kernel level, illustrated in Figure 8. , which is the central part of the Active SLS [10]. The Bridge Ontology consists in different

associations of terms defined in the different ontologies. Some examples are:

```
lmd:Student owl:equivalentClass up:Student ,
                                                    co:Student .
lmd:LearObj owl:equivalentClass lo:LearObj .
```

These triples declare that the Student class from the LMD Ontology is an equivalent class with the Student class, from the User Profile Ontology, and the Student class, from the LO Ontology.

The main properties used in the Bridge Ontology are:

- owl:equivalentClass
- owl:equivalentProperty
- owl:sameAs
- rdfs:subClassOf
- rdfs:subPropertyOf

B. Reasoner, Rules and Querying

The Reasoner allows all queries to be expressed using vocabulary terms from the Application Ontology, even if data is loaded and described through content ontologies. It infers the data with respect to relations declared in the Bridge Ontology. There are two ways in which one can run the Reasoner. The first approach is to make reasoning at the query time (*backward chaining*); this way the Reasoner does the minimum reasoning in order to satisfy the query. The second approach is to load the Reasoner engine and to infer everything possible (*forward chaining*). This way, the Reasoner produces new inferred triples, deduced from the existing data, and generates new relations between the objects defined in the RDF-database. Several Reasoner engines can be used (Pellet [17], Hoolet [20]), each of them supporting different rule languages and different functionalities. For this project, we use the OWLMicroReasoner because it comes with Jena Framework [21], and provides minimum reasoning requirements with fast processing.

Additionally, one can define specific rules in order to establish new relations between the different concepts.

For example, if a LearningObject has its interactivityType set to active, and its interactivityLevel according to the value of the active participation of a learning style that defines a student, then the rule adds an additional relation (lo:matchParticipation) that associates the learning resource to the learning style of that student.

In order to make queries over the RDF-database, we use SPARQL. The Query Constructor builds the queries and forwards them to the Query Executor in order to be performed in the context of the RDF-database. A query example is given below:

```
SELECT ?lo
WHERE {
  ?lo lo:matchParticipation up:Active .
  ?lo app:hasRate ?x
} FILTER (?x>80)
```

The Query Constructor component may save and dynamically combine several queries in order to obtain

more efficient responses. The queries are created with respect to the terminology defined in the Application Ontology.

The application provides an interface in order to create and construct experimental queries over the RDF Datasets. For the normal use of the application, the queries are predefined.

C. Usage Scenario

Different usage scenarios can be envisaged. At registration time, the student completes his learning achievements, acquired skills and diplomas. If the learner wants to take advantage of the personalization functionalities offered by the Active SLS, then he/she is provided with additional information for completing his learning profile. The learning style can be deducted from different tests, and adjusted over time through different key questions.

In the learning process the starting point for the learner is the default path of learning objects from the courses he/she follows. If the learner responds well to lecture, text and reflective participation, then he/she is firstly provided with this type of learning resources. If the student appreciates a learning object as being efficient for his learning, the system can query for other learning resources having the same characteristics. Moreover, the system can propose that learning object to other learners that have same learning profile. The graph is enriched with every interaction the learner has with the system.

Learning objects can be external to the system, but it is necessary for their description to be saved internally. In the default LearnObjectPath, the professor adds all learning objects, and their metadata is verified and registered into the system.

The learner can explore the content by himself, or he/she can take advantage of the system's recommendations, based on: his learning preferences, the resources that he already tagged and rated, preferences from other learners with the same characteristics, etc.

The Semantic Kernel must also provide other functions for the platform. It has to interpret data from the connected modules and dynamically construct queries. The Kernel Interpret is the component that handles the data exchange between the Kernel and User Interface Module.

V. CONCLUSION AND FURTHER WORK

The vision of educational Semantic web-based applications was built starting from the concept of Social Semantic Web, which represents Semantic web applications that employ social software approaches, such as wikis, blogs, etc. Such applications support user collaboration through the use of ontologies, taxonomies or folksonomies. The application of Semantic Web and Social Semantic Web techniques in the domain of eLearning enables a more precise modeling of learners, learning content and context, as well as a more efficient navigation and access to learning material.

This paper presents the concept and the implementation of an Active Semantic Learning System that takes advantage of Semantic Web and Social Semantic techniques to deliver

consistent and pertinent learning resources to the learner, while encouraging learners' active participation through the learning process.

The application is still in a developing phase. The used ontologies must be enriched and refined in order to express data with a high granularity. The interface of the platform has to be intuitive, in natural language and powerful in order to provide access to complex queries. We consider using a more sophisticated reasoning engine and implementing additional rules in order to enhance the power of semantics.

VI. REFERENCES

- [1] Torniai, C. et al. "E-learning meets the Social Semantic Web" Eighth IEEE International Conference on Advanced Learning Technologies, 2008, pp.389-393.
- [2] Jovanovic, J., Gasevic, D., Devedzic, V., "e-Learning and the Social Semantic Web," In Dicheva, D., Mizoguchi, R., Greer, J., (Eds.) *Ontologies and Semantic Web for e-Learning*, IOS Press, Amsterdam, The Netherlands, 2009
- [3] Gruber, T., "Collective Knowledge Systems: Where the Social Web meets the Semantic Web", *Journal of Web Semantics*, 2008
- [4] Felder, R. M., and Silverman, L. K. (1988). *Learning and Teaching Styles in Engineering Education*. Engineering Education
- [5] Balog-Crisan, R., Roxin, I., Szilagyi, I. "Ontologies for a Semantic Quiz Architecture" Ninth IEEE International Conference on Advanced Learning Technologies, 2009
- [6] Sampson, D.G. "Competence-related Metadata for Educational Resources that Support Lifelong Competence Development Programmes". *Educational Technology & Society*, 12(4), pp.149-159.
- [7] IEEE RCD (2004). *IEEE P1484.20.1/Draft Standard for Learning Technology – Data Model for Reusable Competency Definitions*, Retrived January 20, 2010
- [8] IEEE P1484.12.4/D1 Draft Recommended Practice for Expressing IEEE Learning Object Metadata Instances Using the Dublin Core Abstract Model
- [9] S. Bateman, R. Farzan, P. Brusilovsky, G. McCalla, "OATS: The Open Annotation and Tagging System", in the Proceedings of the Third Annual International Scientific Conference of the Learning Object Repository Research Network, Montreal, November 8-10, 2006.
- [10] Szilagyi, I., Balog-Crisan, R., Roxin, I. "Kernel for a Semantic Learning Platform with Adapted Suggestions" Proceedings of the 10th IEEE International Conference on Advanced Learning Technologies, Sousse, Tunisia, 2010, pp.400-402
- [11] <http://www.ocwsearch.com/api/metadata>
- [12] <http://www.dublincore.org/documents/dcmi-terms/>
- [13] <http://www.foaf-project.org/>
- [14] <http://sioc-project.org/>
- [15] <http://www.w3.org/2004/02/skos/>
- [16] http://code.google.com/apis/youtube/2.0/developers_guide_protocol.html
- [17] <http://dbpedia.org/>
- [18] <http://www.holygoat.co.uk/projects/tags/>
- [19] <http://clarkparsia.com/pellet>
- [20] <http://owl.man.ac.uk/hoolet/>
- [21] <http://jena.sourceforge.net/>
- [22] A. Schmidt and C. Kunzmann, "Towards a Human Resource Development Ontology for Combining Competence Management and Technology-Enhanced Workplace Learning," in 1st Workshop on Ontology Content and Evaluation in Enterprise (OntoContent 2006), vol. 4278. Springer, 2006, pp. 1078–1087