UNIVERSITY OF TRENTO

DIPARTIMENTO DI INGEGNERIA E SCIENZA DELL’INFORMAZIONE

38050 Povo – Trento (Italy), Via Sommarive 14
http://www.disi.unitn.it

ONTOGRAPHY SUPPORTED PERSONALIZED ELEARNING REPOSITORIES

Biswanath Dutta, Devika P. Madalli, and A. R. D. Prasad

August 2009

Technical Report # DISI-09-052

Also: the paper is published in "2nd International Conference on the Semantic Web and Digital Libraries (ICSD), Trento, Italy, 2009".
Ontology Supported Personalized eLearning Repositories

Biswanath Dutta¹, Devika P. Madalli², A. R. D. Prasad³

Documentation Research and Training Centre (DRTC), Indian Statistical Institute (ISI), 8th Mile Mysore Road, Bangalore- 560059 (India)
{bisu¹, devika², ard³}@drtc.isibang.ac.in

Abstract. Most of today’s eLearning systems provide static learning materials that are based on the one-size-fits-all philosophy and do not provide a personalized learning space. They are incapable of retrieving and displaying learning materials based upon each individual student’s learning goals. In our present work we identify the core problems that are present in current eLearning systems. Accordingly, we propose possible solutions, upon which we develop a personalized learning system. We deploy a facet based modular structure for this purpose. This system is built upon “semantic learning layer cake”.

Keywords. eLearning, semantic web, ontology, pedagogy, N3Logic rules

1 Introduction

One of the key characteristics of Information and Communication Technology (ICT) is its ability to provide flexible access to information and resources. In learning environment we describe flexible access as access to information and resources at a time, place and pace that are suitable and convenient to individual learners. Through eLearning, where education is delivered online on a mass scale, the goal is to free individual learners from the constraints of traditional residential educational systems where one had to physically attend class lecturers.

Most of today’s eLearning systems provide static learning materials that are based upon the one-size-fits-all philosophy [1] and do not provide a personalized learning space. Essentially they are incapable of generating learning materials dynamically based upon the learner requirements. Ideally eLearning tools should be able to execute complex queries, considering student learning style, student background knowledge, availability of student network connectivity and so on.

In this direction, many researchers are actively working both from academia and corporate worlds. One such significant work carried out by Stojanovc, L., et al. [7] applied semantic web technology to implement an e-Learning scenario. Their work primarily revolves on ontology-based descriptions of content, context and structure of
learning materials. Their work focuses on dynamically creating course structures, using F-Logic as a representation for their ontology. Henze, N., et al. [8] envisioned personal learning services capable of interpreting metadata-annotated learning resource, understanding their annotations with respect to standard ontologies for learning materials, for e.g. LOM or IMS. They investigated ontologies and metadata for three types of resources such as, domain, user and observation. Their work carried a logic-based approach to educational hypermedia using TRIPLE, a rule and query based language for the semantic web. Dongming Xu, et al. [9] implemented a prototyped multiagent-based Personalized Virtual Learning Environment (PVLE) which is based upon the conceptual models of learner, curriculum, situational and pedagogical basis. The research carried out by Verbert, K., et al. [10] is closely related to the content re-usability approach that is applied in our system.

Our work is significantly different from the aforementioned research in that we have placed emphasis on fine-grained description of the learning components. We believe that, to address the two most challenging requirements of the present eLearning system, namely customized/ personalized and dynamic learning spaces, a fine-grained description and contextual representation of the learning space is the solution. The central part of our work is the binding of the educational context in a modular architecture. Equal emphasis is placed on all the three important aspects, such as, content, context and structure.

In our present work we try to identify some core issues of the present eLearning systems and try to determine the root causes of those issues. Accordingly, we propose possible solutions upon which we develop a personalized learning environment. Presented here is ontology supported learning system, where ontologies are expressed in OWL-DL [2] using Protégé ontology editor [3] and developed the logic rules using N3Logic [4].

The following sections are organized as follows, section 2 discusses the eLearning characteristics; section 3 lists the existing problems, issues of the present eLearning systems; section 4 discusses ontology and eLearning; section 5 formalizes the Conceptual Framework of semantic learning space; section 6 deals with the learning ontologies; section 7 presents two use case scenarios of personalized leaning services and section 8 concludes the paper.

## 2 eLearning Characteristics

Before identifying the problems of the present eLearning systems, we list some of the important characteristics required for an ideal eLearning system. They are,

1. Learner centric approach - This approach empowers the learner by facilitating to move away from teacher centered learning systems. Typically, in class room teaching, a teacher decides the agenda and is often only the active participant, whereas students are mostly passive participants.

2. Flexibility (time & space independent) - This adopts the flexi-time approach. A learner with his/ her daily busy schedule, can have flexibility in participating with the learning process. They can adjust the pace of study to other obligations (e.g. family, work, sport). They are not bound to a semester or strict timetable based educational system.
3. Customized and/or personalized content - The learning content is determined by a group of learners or by the individual learner’s based on their needs and aims.
4. Non-linear content - This allows direct access to knowledge in whatever sequence that the learner is looking for, unlike static learning.
5. Continual learning - Learning runs continuously in parallel loops and never stops.
6. Interactive Learning - A common misconception about eLearning is that the absence of human interaction means that there is no one at hand to help learners with their problems. But it reality it works in the reverse order. It facilitates more chances to have someone around the learners (24/7) to help them with their respective problems.
7. Dynamic content - Content is displayed automatically and continuously for a given user based on the users input, experiences, new practices and heuristics.
8. Systematic Learning - Occurs as an integrated activity.
9. Distributed Content - This content is generated from educator-learner interactions.

From the above list, numbers 3 and 7 are missing from most eLearning systems even though they are deemed as most important characteristics.

3 Issues in eLearning environment

It is most important to build and maintain confidence of student community in the eLearning environment. However, there are many problems with the present eLearning systems such as,

1. Lack of group and personalized learning spaces
2. Presentation of the entire learning material as continuous text or media, instead of only relevant information that is actually sought by a learner.
3. Learning sequence
4. Reusability
5. Lack of semantic interoperability
6. Quality assurance
7. Ranking

Some of the above issues may be attributed to the unsystematic organization and description of the learning resources. A personalized learning system can be achieved by fine-grained description of the learning objects. The fine-grained descriptions of learning objects also allow the semantic interoperability. It also allows presentation using different modules as per user requirement. It is often seen that in order to reuse the learning content of a document (e.g., a paragraph, an illustration, a table, etc.), we copy and paste the content into a new document. It is quite possible to reuse the learning objects in a more sophisticated way if we can access the specific components of a learning object and re-purpose them on-the-fly [6].

Specifically, we can formalize an eLearning system into three different levels, such as, content, context and structure, where, content identifies what the learning material is about, context identifies in which form this topic is presented, while structure is to
4 Biswanath Dutta1, Devika P. Madalli2, A. R. D. Prasad3

comprehend the arrangement of the learning materials in a learning course [7]. In our present work, we tried to address the following issues:

1. Learner centric educational architecture
2. Interest based knowledge retrieval
3. Achieving semantic interoperability
4. Achieving reusability

4 Ontology and eLearning

In order to describe the learning objects, different communities use different metadata standards as per their requirements. We know that metadata elements lack a formal semantics as they are mainly useful in indexing the documents. So when it comes to sharing resources between heterogeneous domains instead of homogenous domains, we face the problem of incompatibility. The lack of shared understanding between terms of various metadata vocabularies might be avoided by using ontologies as a conceptual backbone in an eLearning scenario [7]. Ontology in general helps us to define the learning components more strictly. According to the most popular definition of ontology (as defined by Gruber) is “an explicit specification of a conceptualization” [11]. The purpose of ontology is to favour interoperability by providing a common terminology and understanding of a given domain of interest, which in turn allows for the assignment of a clear meaning to learning materials. Our system is implemented in the semantic web environment, and emphasizes the fact that:

1. Standards are concerned about semantics rather than just about syntax
2. Extensible methods for data integration should be provided in eLearning environment

5 Conceptual Framework of Semantic eLearning

In order to deal with the above mentioned problems, we formalize a conceptual learning space and call it a semantic learning layer cake. Each layer is built on top of another layer. These layers are formed in a bottom-up approach [figure 1].

Fig. 1. Semantic learning layer cake
The bottom layer contains the content objects. The set of content objects form the learning object. Figure 2 shows the internal structure of the learning objects. We perceive learning objects as an aggregation of content objects whereas each content object is formed by a set of content fragments (or content units). The content fragments are learning content elements in their most basic form, like text, image (e.g., paintings, graphics, moving images), sound, datasets (e.g., tables, lists), etc. They represent individual resources in isolation. We also see from the figure that the aggregated content fragments define the navigation within a content object which helps us in defining the learning path.

In the semantic learning layer, on top of learning object is the content and context. The content contains the concepts or subject terms defining the “thought content” or “semantic content” of the learning objects. In our framework, we formalize content as “domain ontology”. “Domain ontology” helps in reducing the knowledge gap between the teacher and learner by formulating the unambiguous and shared understanding of terms. It also helps to overcome the problem of synonym, homonym, antonym, etc. and other related problems (e.g., acronym) that we often face in an online information retrieval environment. The context identifies the facts or circumstances of the learning objects. In our framework, the context is represented from three different angles, such as, matching the education level of both the document and the learner, intended use of the learning object and the learning objectives. The details are provided in the following sections.

On top of content and context is the structure layer. Structure formalizes the relations between the learning materials. The relations are specified by the properties, such as, hasPart, isPartOf, hasPrerequisite and isPrerequisiteOf. These relations help in defining the learning sequence. Learning design is the top most layer of the semantic learning cake. This layer uses the students profile and other layers below it to create a personalized learning environment with the aid of sequential activities. Essentially, it specifies the roles, sequence, logistical information and pedagogical information. The learning design layer is formalized using logic rules.

6 Ontology Framework

In this section we present the ontology frameworks developed to create the personalized learning environment. Our developed ontologies are modular based. The
advantages of modular based ontology are, it is easy to manage and easy to incorporate the new set of concepts within the ontology at any point of time. The main components of our ontology are, document ontology, domain ontology and student ontology. The following figure 3 shows the partial building blocks of our document ontology and figure 6 shows the student ontology framework. The ontologies are developed using the OWL-DL [2]. In deploying the domain ontology (not discussed here as it is outside the scope of this paper) we used SKOS [16] and bounded with OWL-DL.

In order to build our system we created content space containing the learning materials in XML format. In order to present the learning materials to each student based upon the individual learning goals and their profile, it is proposed to used XSLT to compose and transform the XML document into web compatible format. It is important to note here that in order to describe the learning materials, IEEE LOM1 and Dublin Core2 metadata standards are deployed while to create the student profile we used vCard3 and IMS LIP4.

6.1 Document Ontology

Figure 3 shows the partial class (main building blocks) diagram of our document ontology. In the diagram the italic classes represent the abstract classes whereas the non-italic classes are the concrete classes. The class Entity has four subclasses: Program, Course, LearningObject and Contribute. However, in this paper we focus on only those classes which are directly involved in generating the personalized learning services. The class LearningObject describes learning resources, in our case digital and web-based materials that can be used and re-used to support learning. For example, multimedia content, instructional content, learning objectives, instructional software and software tools, persons and organizations, or events referenced during technology supported learning.

The class, LearningObject consists of three subclasses, such as, Topic, ContentUnit and SupportResource. Class ContentUnit is used to describe the item specific information i.e., item level description, for example, lom-tech:location (of the learning objects), requireNCSpeed (require network connection speed), lom-tech:format, hasVersion etc.

---

2  http://dublincore.org/documents/dces/
3  http://www.imc.org/pdi/vcardoverview.html
4  http://www.imsglobal.org/profiles/
Class **ContentUnit** is used to describe the content fragments. The properties of **ContentUnit** are, **hasLearningObjective**, **isVersionOf**, **hasVersion**, **isContentUnitOf**, **lom-tech:format**, **lom-tech:location** and **requireNCSpeed**. It is worth noting that, **hasVersion** holds the information of a content unit available in different formats (e.g., a textual content may have other forms, such as, audio, graphic, etc.). In order to uniquely identify the different forms of content units, we divide the **ContentUnit** class into five subclasses (Figure 4), such as, **Text** (e.g., textual content of the materials, such as, letters, dissertations, poems, newspapers, articles, archives of mailing lists), **Dataset** (e.g., lists, tables, databases), **Image** (e.g., images and photographs of physical objects, paintings, prints, drawings, other images and graphics, animations and moving pictures, film, diagrams, maps, musical notation), **Software** (e.g., Java/C source files, MS-Windows, executables, or Perl script) and **Sound** (e.g., an audio compact disc, and recorded speech or sounds). The class **Dataset** is further divided into three, **Databases**, **Lists** and **Tables**. The **Image** class is divided into **MovingImage** and **StillImage**. **MovingImage** is further divided into three subclasses, **Animations**, **TelevisionPrograms** and **Videos**. **StillImage** is divided into five subclasses, **Drawings**, **GraphicDesigns**, **Maps**, **Paintings** and **Plans**. The class **Sound** is divided into two, such as, **Music** and **Speech**.
Among the other significant classes and subclasses of document ontology [figure 5] are lom-edu:Context, LearningObjectives and lom-edu:LearningResourceType. These bring the learning context into the learning space. The class lom-edu:Context identifies to whom the resource is intended for or useful. For example, the learning materials could be useful for higher education, such as by graduate and undergraduate students or for elementary education, etc. The class LearningObjective represents the educational objectives of the resources, for instance, define, evaluate, introduce, example, compare, classify, demonstrate, describe, design, etc. We have identified a total of 19 concepts that express the educational objectives of the learning objects. In this regard we used Bloom’s taxonomy of educational objectives. [13]. The class lom-edu:LearningResourceType identifies the potential educational use(s) of content associated with the learning resource. It is divided into two subclasses, such as, EducationalResource (e.g., example, exercise, index, lecture, etc.) and ExaminationResource (e.g., exam, project_task, questionnaire and self-assessment).
Class *lom-edu:Difficulty* identifies how hard the learning material is to work with. We identified five difficulty levels, namely, *easy*, *very easy*, *medium*, *difficult* and *very difficult*. Class *lom-tech:Format* assigns the technical datatype(s) of (all the components of) the learning objects. It is divided into two subclasses, namely, *Continuous* (e.g., audio, video) and *Discrete* (e.g., text, image, application).

### 6.2 Student Ontology

Figure 6 shows the class representation of “student” ontology. Here the main class is *Student*, around which the other classes are developed. The classes *Topic* and *Course* are shown in the following figure and are not part of the Student ontology but are part of the document ontology. To define the student knowledge we share the resources of class *Topic* of the document ontology. The property *hasKnowledgeOn* identifies the student background knowledge. On the other hand, to specify which course the student is registered for, we use the *Course* class of the document ontology. The property *registeredForCourse* creates a semantic relation between the classes *Student* and *Course*.

![Class diagram of student ontology](image)

The class *CognitiveLearningStyle* defines cognitive characteristics that the student possesses (discussed in details in section 8). On the other hand each student has unique approach to learning. There are students who prefer to learn by reading materials while some students prefer to learn through graphics or some by just listening and so on. The class *StudentType* captures the variations in student learning styles (discussed in details in section 7).

The class such as, *ims-lip:QCL* identifies the Qualifications, Certifications and Licenses (qcl) of a student. The property *degreeGrantedBy* (identifies the organization that offered the degree) associates the *ims-lip:QCL* class to the
Organization class. The class EducationLevel identifies the student level of learning. Interest class identifies the student interests. The Language class identifies the student languages in three different levels, expressed by the properties, languageProficiency, languageDisability and languageEfficiency. The Competency class identifies the student competency or skill level. The class Time identifies the temporal information and is divided into four subclasses, namely, Day, Month, Season and Year.

We adopted a faceted approach [14] in design. For example, the classes, such as Language and Time are treated as the common modifiers applicable to document ontology, students ontology and domain ontology.

7 Use Case Scenario

Understanding the learners and their characteristics is an important aspect in any mode of the teaching learning process, be it a personal class room teaching session, distance mode of learning or eLearning space. In class room teaching it is quite easy for a teacher to understand the student behavior and accordingly a teacher can deliver the best suitable material(s) to meet the individual learner characteristics. Since the eLearning space is a self-learning space, we need to design the system with care keeping in mind the similar space as we have in the face-to-face class room teaching. The aim is here to make the learning process sensitized to the individual learner. Here we discuss two such significant characteristics of learners which need to be taken care in developing personalized eLearning system. The first one is “Learning style” and the other one is “Cognitive style of learning”.

Case 1: Learning style

Each student has his/ her own way of learning. There are students who prefer to learn by reading materials while some students prefer to learn through graphics or some by just listening and so on. To identify the varieties we use Fleming’s VARK model, the most popular and relatively simple model that covers all the aspects of a learner’s learning styles [18]. The model characterizes the learners into four categories that expand upon Neuro-linguistic programming (VAKOG or known as the 4-tuple [19]) models:
1. visual learners - have a preference for seeing (think in pictures; visual aids such as slides, diagrams, handouts, etc.)
2. auditory learners - best learn through listening (lectures, discussions, tapes, etc.)
3. reading/writing learners – prefer to learn by reading/ writing the textual materials
4. kinesthetic learners or tactile learners - prefer to learn via experience— moving, touching, and doing (active exploration of the world; science projects; experiments, etc.).

In order to provide the documents based upon the student type (learning style), we created a class, called StudentType in our “student ontology”. Four instances are created for the class StudentType, such as, visual; auditory; reading-writing and kinesthetic. In our document ontology, we have a class, called, lom-tech:Format which is further divided into two, namely, Continuous and Discrete. For the class,
Continuous we defined two instances, such as, audio and video and for the Discrete class we have defined instances, such as, text, application and image. The format is assigned for each content unit for the topics. lom-tech:format property is assigned for the stated purpose where the domain and range of it is ContentUnit and lom-tech:Format respectively.

To generate personalized services, we match the student type with the document format. For example, if a user has reading-writing habit, then it is understood that s/he learns through the textual materials. Similarly, the auditory learners prefer to “read” through the audio materials. In the same way, visual learners prefer to have the video, image materials and kinesthetic learners prefer to have the application related materials. The rules are written following the N3Logic rules, a subset of First Order Logic (FOL). As N3Logic is a subset of the FOL, it is more expressive and is useful as a tool in the open Web environment. [4]. Rules are built to reason over distributed information sources (ontologies).

\[
\text{@forAll } S, D, C, V. \\
\{S \text{ a stu-onto:Student; stu-onto:typeOfStudent stu-onto:reading-writing. C lom-tech:format text} \} \Rightarrow \{S :eligible_to_receive C\}. \\
\]

Here, S is a type of student and has reading-writing habit. According to the above discussion, here the condition is, if a student has reading-writing habit, s/he must be given the textual documents only. We see here that the learning items (C) with textual format should be provided to the reading-writing learners. In order to meet the individual learning goal the learner search through subject term. During the search time the system takes care of the individual student learning style as it recognizes the learner once the learner logs into the system. The query could be formulated like the following. In the query we call the conclusion of the above rule as premises.

\[
\text{@forAll } S, C, D, V. \\
\{D \text{ dc:subject dom:Pre-coordinate indexing; hasContentUnit C. C hasLearningObjective define. S :eligible_to_receive C} \} \Rightarrow \{S :receive_with C\}. \\
\]

It is to be noted here that, in order to meet the context based search, the learner can set the learning context from the interface during the search time. As we see here, the learning objective is “define”. Similarly the learners can assign the other learning context (discussed in section 6.1) during the search time. The following rules shows the rules for other three types of learners.

\[
\text{\{S a stu-onto:Student; stu-onto:typeOfStudent stu-onto:auditory. C lom-tech:format audio\} \Rightarrow \{S :eligible_to_receive C\}.} \\
\text{\{S a stu-onto:Student; stu-onto:typeOfStudent stu-onto:visual. C lom-tech:format video\} \Rightarrow \{S :eligible_to_receive C\}}. \\
\]
Case 2: Cognitive learning style
Cognitive learning style (CLS) defines cognitive characteristics that the student possesses. These can be inferred from cognitive tests as the Ross and Witkin tests [17]. The cognitive characteristics formalize the type of information processing and reasoning the student uses. These properties are useful in user modeling, so that contents can be tailored for each student’s characteristics. We identified five instances as student cognitive learning style, such as, Analogue-Analytic, Concrete-Generic, Deductive-Evaluative, Relational-Synthetic and Indefinite style.

In order to model CLS, in the student ontology we have a class called, stu-onto:CognitiveLearningStyle and defined five instances of it as discussed above. Now consider a case, for example, a student with deductive-evaluative style of learning has learning goal “pre-coordinate indexing”. This type of student prefers to learn first “pre-coordinate indexing” by taking the theory first and then go for practice. Whereas the concrete-generic type of students with the same learning goal prefer to learn by taking the examples first and then figure out the theory. In order to meet the above goal we created a class called lom-edu:LearningResourceType which is further divided into two, such as, EducationalResource and ExaminationResource. Under the EducationalResource we have the instances, example, exercise, lecture, index, problem_statement, simulation and tutorial; whereas, for ExaminationResource we have the instances, namely, exam, project_task, questionnaire and self-assessment.

In the following rule example, we show the rule made for delivering the learning materials to fulfill the learning goal on “pre-coordinate indexing” suiting the individual learner CLS. Here we consider a student with “deductive-evaluative” learning style. The condition is, s/he must receive the lecture materials first followed by exercise materials.

@forAll S, D, scp, C, L1, L2, L3, sub.
(S a stu-onto:Student; stu-onto:hasCognitiveLearningStyle stu-onto:Deductive-Evaluative. (scp 2) e:findall (C {D dc:subject dom:Pre-coordinate_indexing. D lom-edu:learningResourceType lecture. D hasContentUnit C} L1). (scp 2) e:findall (C {D dc:subject dom:Pre-coordinate_indexing. D lom-edu:learningResourceType exercise. D hasContentUnit C} L2). (L1 L2) list:append L3 => (S :WORKING_deductive_student_get_list L3).

Here, S is a student with deductive-evaluative learning style. D is a topic with content unit C (for details see section 6.1). L1 is a list of lecture materials where as L2 is a list of exercise materials. L3 is a lists containing the lists L1 and L2 maintaining the order L1 followed by L2.
It is worth to mention here that, we use built-in property, “findall” (e:findall) from Euler [20], an inference engine supporting logic based proofs. The syntax of e:findall is as follows,

```prolog
```

It unifies ?ANSWER with a list that contains all the instantiations of ?SELECT satisfying the ?WHERE clause in the ?SCOPE ?SPAN of all asserted n3 formulae and their log:conclusion.

In order to avail the personalized services we can also consider the student education level, competency level, language proficiency, etc. in matching with the document ontology components.

8 Conclusion

In this paper we presented our ongoing work on developing a personalized learning space. The above discussion shows that if we describe the learning space in a more meaningful and fine grained manner, it is possible to solve most of the existing problems of the present eLearning systems, such as, re-usability, provisions for delivering specific information as per the student’s learning goal instead of the entire learning material, composition and sequencing of the learning contents. The area of semantic operations in digital repositories is akin to the demonstrated system discussed here.

9 References

16. Introduction to SKOS. http://www.w3.org/2004/02/skos/intro