

## On-Demand Construction of Personalized Learning Experiences Using Semantic Web and Web 2.0 Techniques

Nicola Capuano, Matteo Gaeta, Francesco Orciuoli, Pierluigi Ritrovato  
*Dipartimento di Ingegneria dell'Informazione e Matematica Applicata - University of Salerno*  
*Via Ponte don Melillo, 84084, Fisciano (SA), Italy*  
*{capuano, gaeta, orciuoli, ritrovato}@diima.unisa.it*

### Abstract

*Nowadays, the Semantic Web technologies are exploited also in the e-learning domain in order to provide personalized and adaptive learning experiences, semantic annotation of learning contents and learner profiling. The approaches of the Web 2.0, instead, are used to implement and deploy knowledge exchange services based on the concept of social collaboration. In this work, we propose an approach resulted from the convergence between Semantic Web and Social Web to manage, agilely and easily, the contingent learning needs of workers within organizations. Our intention is to support the use of natural languages to express the learning needs for either driving the automatic generation of learning units or effectively adapting learning pathways.*

### 1. Introduction

Inside the organizations, the Technology Enhanced Learning (TEL) solutions still suffer from a content-centered vision and approaches of Learning Management Systems (LMS) and Learning Content Management Systems (LCMS) with very limited personalization capabilities, whose pedagogies are embedded in the content itself. In an organizational environment the TEL solutions are easily integrated with general learning objectives through the setting of pre-determined learning courses for class of employees according to their learning pathways. Moreover, even as we speak, the adoption of a Web 2.0 approach emphasizes the central role of the user and, more in general, the user-generated content, ratifying the shift from Information Society to Knowledge Society. Indeed, completely missing, especially in the working environments, are just-in-time learning services allowing the users to ask for (organize) dynamically generated and personalized e-learning experiences that can meet their specific needs but also aligned to the

organizational objectives, and for an agile mechanism to harvest the collective knowledge, sharable with peers, enriching the learning process. With respect to the previously described scenario, we will focus on the two following problems: (a) Separation and missing interconnection between organization learning pathways and real contingent workers' needs. Usually, the organization human resource management departments are very slow in detecting the workers' training needs and in organizing learning activities basing on such needs, so when the learning actions start they are often out-of-date or address very general objectives; (b) Incapacity to exploit the collective knowledge, within the organisation, in order to satisfy the workers' learning needs. A learning objective, though not inserted in organization learning pathways, could be already achieved by a more experienced worker in the same organization.

In this paper we propose an integrated approach able to fill the gap between organization learning paths and real workers' needs. This is achieved by exploiting an innovative Learning Model, which enables both the definition and delivery of personalized e-learning experiences (based on the main principles of the Semantic Web) and a natural language human-computer interface able to support the users in defining their learning needs. The proposed approach exploits the knowledge already acquired by an organization workers in order to support other workers to fulfill their own learning needs applying the model of the so-called Collective Knowledge Systems (CKS) in the e-learning domain ..

The paper is organized as follows: Section 2 describes the state of the art of systems and approaches that make use of knowledge and semantic technologies for learning; Section 3 presents our approach to the creation of "just-in-time" learning experiences according to user's needs expressed in a natural language; in Section 4 we propose an approach based on Collective Knowledge Systems to cover the "lack of

knowledge” the created learning experiences have left out ; Section 5 draws the conclusion and future works.

## 2. Background and related works

The use of Semantic Web techniques, to effectively organize and manage available e-learning resources according to peculiar necessities of both teachers and students, has been advocated by many authors. Brase and Nejdil in [1] have showed the increasing importance, in the e-Learning field, of knowledge modeling through metadata definition standards.

However, these standards introduce the problem of incompatibility between disparate and heterogeneous metadata descriptions across domains, which might be avoided by using ontologies as a conceptual backbone in an e-learning scenario [2]. A number of systems have been developed to handle learning resources by means of Semantic Web technologies. Edutella [3] is an open source project for P2P network users interested in the exchange of learning resources. Another representative system that uses Semantic Web techniques in an e-Learning environment is the Courseware WatchDog [4]. Furthermore, a great number of very attracting services are rising from the Web 2.0 and offer abundance of knowledge sharing [6]. The headliners of these services are Wikipedia, MySpace, YouTube, Flickr, Del.icio.us, Facebook and Technorati. Particular types of knowledge exchange services are Yedda (<http://yedda.com>) and Yahoo Answers (<http://answers.yahoo.com/>) that people can use to ask questions which all can respond to, sharing knowledge about the given subject. The CKSs, whose basic principles lay upon the features of systems like Yedda, are defined in [5] as software that can deliver the opportunity of collective intelligence. The CKSs enable the collection, harvesting and sharing of large amounts of human-generated knowledge supported by machines. In a typical CKS, obtained by the convergence of Social Web (class of Web 2.0 sites and applications in which the participation of the user is the primary driver of value) and Semantic Web techniques, the main actors are: (a) a community of contributors participating in a social process augmented by computer mediated communication and long-term memory, (b) a search engine for retrieving information, (c) intelligent users who actively query the system using strategies tuned to content generation process and search engine.

Actually, at present no CKS is focused on the learning needs sharing to influence the learning planning and to enrich the learning resources repositories (i.e. storage mechanism for artifacts used in the setting and execution of learning processes).

## 3. Needs-driven e-learning experiences

In this section we present an approach to define personalized e-learning experiences based on people’s learning needs expressed in a natural language. The building process can be carried out by: (i) a software module that assembles the Unit of Learning (e.g. a package including all the information useful to the Learning Management System in order to deliver the personalized e-learning experience) starting from an ontology [7] modeling the educational domain (e.g. Mathematics, Physics, etc.); (ii) a learner profile indicating cognitive state and learning preferences of a specific learner; (iii) a digital repository where learning objects are annotated with standard metadata schemas; and (iv) a set of course parameters specifying learning context, pre-requisites and other attributes.

### 3.1. Personalized UoL building process

The foundation for the Unit of Learning (UoL) building process is the *Learning Model* described in [8]. The *Learning Model* allows to automatically generate a UoL (i.e. a course, a module or a lesson structured as a sequence of Learning Activities represented by Learning Objects and/or Learning Services) and to dynamically personalise it during the learning process according to the learners’ preferences and cognitive state. In order to achieve the expected adaptation capability, the *Learning Model* uses three specific sub-models: *Knowledge Model*, *Learner Model* and *Didactic Model*.

The *Knowledge Model* describes, in a machine-understandable way, the subset of the educational domain that is relevant to the e-learning experience we want to define, concretize and broadcast. The *Knowledge Model* exploits the ontologies, i.e. engineering artifacts, constituted by a specific vocabulary used to describe domains, plus a set of explicit assumptions regarding the intended meaning of the words from the vocabulary. In our approach the vocabularies are composed of terms representing subjects that are relevant to the educational domain we want to model. Some subjects are associated to other subjects through a set of different conceptual relations. The most important relations are: *HasPart* that is a part-of relation and *IsRequiredBy* that is an order relation. The ontologies constructed following the few aforementioned informal rules are called e-learning ontologies. Observe that the concepts in the e-learning ontologies we are referring to are the subjects of the educational domain we are modeling. Let us now illustrate how to build an e-learning ontology. Suppose

we have to model the educational domain  $D$ , so we try to conceptualize the knowledge underlying  $D$  and find a set of terms representing relevant concepts in it. The result of the previous step is the list of terms  $T = C, C_1, C_2, C_3$  where  $T$  is one of the plausible conceptualizations of  $D$ . The existence of the relations *HasPart*( $C, C_1$ ), *HasPart*( $C, C_2$ ) and *HasPart*( $C, C_3$ ) means that in order to learn a subject  $C$  the learners have to learn subjects  $C_1, C_2$  and  $C_3$  without considering a specific order. If we add the relations *IsRequiredBy*( $C_1, C_2$ ) and *IsRequiredBy*( $C_3, C_2$ ) to the previous set of relations we can state that  $C_1$  has to be necessarily learned before  $C_2$  and  $C_3$  has to be necessarily learned before  $C_2$ . The learner is the main actor of the whole cognitive process and he/she has to be represented in a machine-understandable way in order to support the personalization process. The *Learner Model* states that the learner is represented by a cognitive state and a set of learning preferences. For the sake of simplicity, we disregard the *Didactic Model* being not essential for the focus of this work.

The *Learning Model* allows the construction of personalized e-learning experiences through the execution of a building process based on some specific algorithms. Using these algorithms it is possible to generate courses tailored to a class, to a specific group and even to single learners. In the following, we present the steps needed for the creation of an e-learning experience tailored to a single learner's cognitive state and preferences. In our approach, a Learning Object is a learning content (or a packaged aggregation of learning contents) that can be delivered through a Web Browser, that is annotated with an instance of a metadata schema interoperable with IEEE LOM<sup>1</sup> and that is stored and indexed into a Learning Object Repository. Exploiting some API exposed by the Repository we can get, for instance, all the Learning Objects whose metadata (in the field "description") contain the word *mathematics*. Furthermore, a single Learning Object can be associated to an e-learning ontology through a metadata field, namely "subjects list", that stores the references to all subjects, coming from one or more ontologies, explained by the Learning Object content.

The e-learning experiences are defined as: (i) a set of *Target Concepts* ( $TC$ ), i.e. the set of high-level concepts to be transmitted to the learner; (ii) A *Learning Path* ( $LP$ ), i.e. an ordered sequence of atomic concepts (subjects) that is necessary to explain to a learner in order to let him/her learn  $TCs$ . Given the personalization on a particular learner, the sequence does not contain subjects already "learnt" by that learner; (iii) A *Presentation* ( $PR$ ), i.e. an ordered list of

Learning Objects that the learner has to use in order to acquire knowledge about subjects included in  $LPs$ .

The  $LP$  can be automatically obtained basing on  $TCs$ , whilst a  $PR$  can be automatically constructed from the  $LP$  and querying (using metadata information) one or more Learning Object Repositories. The  $TC$  can be settled by the teacher or the learner directly (in case of self-directed learning) and can be obtained by manually selecting concepts on the e-learning ontologies. Excluding the selection of  $TCs$  and other customization parameters, the building process is fully automatic and realized through the execution of several algorithms. The most important are: *Learning Path Generation Algorithm* and *Presentation Generation Algorithm* [9]. Once the *Presentation* (for a given student) is ready, it is packaged into a UoL, then the enrolled student can start his/her execution phase.

### 3.2. Extending the Learning Model with Learning Objectives support

The MIT Teaching and Learning Laboratory (<http://web.mit.edu/tll/teaching-materials/learning-objectives/index.html>) states that a Learning Objective should describe what students should know or be able to do at the end of the course that they couldn't do before. A good example of a Learning Objective could be: *Solve problems using matrix techniques and algorithms*. The experts in the act of writing Learning Objectives often talk about using the acronym **SKA** that stands for: **Skills** (i.e. what students should be able to do by the time the course is completed), **Knowledge** (i.e. what students should know and understand by the time the course is completed) and **Attitudes** (i.e. what the students' opinions will be about the subject matter of the course by the time it is completed).

With respect to the *Learning Model*, and in particular with the *Knowledge Model*, a Learning Objective can be defined as: (**Title**,  $C_1, C_2, \dots, C_n$ ). **Title** is a text representing the Learning Objective in a natural language,  $C_i$  can be a skill, a piece of knowledge or an attitude required by the objective. Using the e-learning ontologies we can simply model Skills, Knowledges and Attitudes for a given educational domain [10]. We can choose to have, for a specific educational domain, three e-learning ontologies respectively, modeling Skills, Knowledges and Attitudes (this will be our assumption for the following sections) or, alternatively, a single ontology. So,  $C_1, \dots, C_n$  are only references to Skills, Knowledges and Attitudes ontologies. Instances of Learning Objectives, as we have previously defined, can be created, stored and indexed independently from the definition of a personalized e-learning experience in

<sup>1</sup> <http://ltsc.ieee.org/wg12>

order to settle a searchable Learning Objective Repository.

The *Learning Model* can be easily extended substituting the set of *Target Concepts* (used by the instructors to indicate the set of high-level concepts the learner can gain, through the delivery of a personalized e-learning experience) with a set of Learning Objectives selected from the repository or created on the fly. The UoL building is exactly the same as that illustrated in section 2.2. The only difference is the two main algorithms (*Learning Path Generation Algorithm* and *Presentation Generation Algorithm*) need to be executed for each different ontology referred within the experience definition.

### 3.3. Learning Needs as a high level access to self-learning personalized experiences

During the execution of project activities, the workers could feel the lack of expertise on one or more topics related to the project they are involved in. Obviously, they can express their needs easily and quickly using a natural language rather than navigating catalogues of learning resources, and personally assembling suitable learning units that satisfy their educational needs.

We think that providing a high-level access interface able to educational requests (expressed in a natural language) coming from workers and automatically assembling a personalized learning experience aimed at satisfying their needs, can considerably reduce the delay between the time when a worker understands the need to learn and the time when he/she can start his/her learning process.

In our approach, a software component processes the natural language sentence, discovers and assembles the best available learning resources considering organization, group and individual learning objectives modeled and represented into the system. The discovery process can take benefits from the exploitation of techniques able to find a similarity value between the sentence representing a worker Learning Needs and the textual annotation of Learning Objectives, as we have defined in section 3.2.

We think that *Sentence Similarity* approaches could be contextualized and applied in order to define models suitable to support high-level access functions to personalized e-learning experiences. In [11], the authors present an interesting approach to *Sentence Similarity* based on an algorithm that takes account of semantic and words order information in very short texts. The aforementioned algorithm computes the semantic similarity between two sentences using information from a structured lexical database (i.e.

WordNet), from the application of text algorithms for *stemming*, *string-similarity* (e.g. Levenshtein Distance, etc.) and from corpus statistics. The use of a lexical database enables our method to model the human common sense knowledge, while the incorporation of corpus statistics allows the method to be adaptable to different domains.

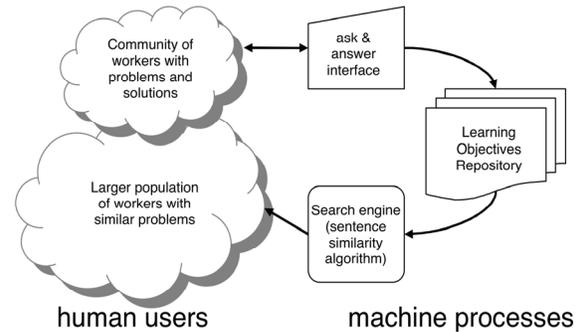


Figure 1. Functional View

Our approach is focused on matching Learning Needs expressed in a natural language by workers, through a user-friendly interface, with Learning Objectives defined within the Learning Objective Repository. The matching operation consists in executing the *Sentence Similarity Algorithm* (it is defined exploiting and contextualizing to the e-learning the algorithm presented in [14]) between the text representing the expressed Learning Need ( $LN$ ) and the texts of all the Learning Objectives ( $LO_i$ ) stored into the repository. The result of the algorithm execution is the list of all sentence similarity measures  $S_i = S(LN, LO_i)$ . Only the objectives  $LO_i$ , such that  $S_i$  is bigger than a given threshold, can be presented to the worker who can select one (or more than one) of them and request the delivery of a personalized e-learning experience (with respect to the modalities indicated in section 3.1 and 3.2) in order to meet his/her needs. The definition of a personalized e-learning experience starts from one or more Learning Objectives and can adapt the *Learning Path* and the *Presentation* using information stored inside the learner profile..

## 4. Exploiting learning needs sharing to enrich e-learning repositories

The approach proposed in section 3.3 can be also extended to exploit unsatisfied Learning Needs in order to enrich the learning resources repositories.

The knowledge is not static but evolves over time. The rapid evolution of knowledge can cause several problems within an organization. One of these problems is just the population of the learning

resources repositories. We think that the idea of collective intelligence (within the Web 2.0 communities) can be exploited in order to solve the aforementioned problem. In our approach we aim at applying the models underlying the CKSs to exploit the organisation collective intelligence and to keep up-to-date the learning resources repositories, using content provided by the workers community, starting from their Learning Needs. In Figure 1, the functional view of the system is reported. The workers provide their needs to the system through an “ask & answer interface” supporting natural languages. A “search engine” improved with a “sentence similarity algorithm” (presented in section 3.3) is used to find relevant Learning Objectives already defined and stored in the repository and to present the results to the users (“Larger population of workers with similar problems”). Not all the expressed educational requests (Learning Needs indicated by the high level access interface) can be satisfied by the “search engine”, cause a lack of suitable Learning Objectives. So, the unsatisfied requests could be spread upon the community through the use of Social Web techniques and, in particular, by the “ask & answer interface” to the “Community of workers with problems and solutions”. Once the community receives the request, the workers (counting on suitable expertise) have the opportunity to produce user-generated learning resources (i.e. e-Learning Ontologies, Learning Object and Learning Objectives) able to populate the repositories and to directly or indirectly satisfy the request. A notification about the new presence of Learning Objectives within the repository related to a given Learning Need is sent (also by e-mails) to the workers who expressed the unsatisfied needs.

## 5. Conclusions and future works

In this work we have proposed our approach to solve two problems related to the self-learning processes within the organisations: the separation between real workers’ learning needs and learning pathways provided by typical TEL solutions and the incapacity to exploit a wide knowledge source provided by the workers’ collective intelligence in order to improve the organisation learning processes, enabling the continuous population of learning resources repositories and to increase the quality of learning planning. With respect to the first problem we have proposed the integration of a semantic-based *Learning Model* with an algorithm for the sentence similarity computation, in order to provide with the automatic definition and delivery of personalized e-learning experiences considering learning needs

expressed in a natural language. As for the second problem instead, we have proposed an approach based on the Collective Knowledge System model in order to share, within a community (workers), the Learning Needs which are unsatisfied by the existing defined Learning Objectives. The Learning Needs sharing will produce two positive effects: (i) the possibility for the workers community to provide suitable learning resources and satisfy specific Learning Needs, and (ii) the opportunity for the organisation to adapt learning plans on the basis of real workers’ learning needs. A first implementation of the proposed integrated approach has been provided extending the capabilities of a commercial e-Learning Platform, namely Intelligent Web Teacher (IWT) [12]. The sentence similarity computation feature has been realized using Lucene.NET [13] full-text search engine.

We aim in the future at extending IWT by implementing the CKS for the Learning Needs proposed in section 4.

## 6. References

- [1] J. Brase and W. Nejdl, “Ontologies and Metadata for eLearning”, *Handbook on Ontologies*, 2004, pp. 555-574.
- [2] L. Stojanovic, S. Staab and R. Studer, “eLearning based on the Semantic Web”, *WWW and the Internet*, 2001.
- [3] W. Nejdl, B. Wolf, S. Staab and J. Tane, “EDUTELLA: Searching and Annotating Resources within an RDF-based P2P Network”, *Semantic Web Workshop*, 2002.
- [4] J. Tane, C. Schmitz, G. Stumme, S. Staab and R. Studer, “The Courseware Watchdog: an Ontology-based tool for finding and organizing learning material”, *Mobiles Lernen und Forschen*, 2003, pp. 93-104.
- [5] T. Gruber, “Collective Knowledge Systems: Where the Social Web meets the Semantic Web”.
- [6] J. P. Allen, “How Web 2.0 Communities Solve the Knowledge Sharing Problem”, *ISTAS 2008*, pp. 1-3.
- [7] T. R. Gruber, “Towards Principles for the Design of Ontologies Used for Knowledge Sharing”, 1993.
- [8] G. Albano, M. Gaeta, S. Salerno, “E-learning: a model and process proposal”, *IJKL*, vol. 2, no.1, 2006, pp. 73–88.
- [9] N. Capuano, M. Gaeta, S. Miranda, F. Orciuoli and P. Ritrovato, “LIA: An Intelligent Advisor for e-Learning”, 1st world summit on The Knowledge Society, 2008.
- [10] <http://www.nwlink.com/~Donclark/hrd/bloom.html>
- [11] Y. Li, D. McLean, Z. Bandar, J. O’Shea and K. A. Crockett, “Sentence Similarity Based on Semantic Nets and Corpus Statistics”, *IEEE TKDE*, 18(8), 2006, pp.1138-1150.
- [12] G. Albano, M. Gaeta and P. Ritrovato, “IWT: an innovative solution for AGS e-Learning model”, *IJKL*, vol. 3, n. 2-3, 2007.
- [13] <http://incubator.apache.org/projects/lucene.net.html>