CADDIE and IWT: two different ontology-based approaches to Anytime, Anywhere and Anybody Learning

Giovanni Adorni¹, Serena Battigelli¹, Diego Brondo¹, Nicola Capuano², Mauro Coccoli¹, Sergio Miranda², Francesco Orciuoli², Lidia Stanganelli¹, Angela Maria Sugliano¹, Giuliano Vivanet¹

¹DIST – Department of Communication Computer and System Sciences, University of Genoa, ²DIIMA - Department of Information Engineering, and Applied Mathematics, University of Salerno

Abstract

The Semantic Web seems to offer great opportunities for educational systems aiming to accomplish the AAAL: Anytime, Anywhere, Anybody Learning. In this scenario, two different research projects are here introduced: CADDIE (Content Automated Design & Development Integrated Editor), developed at the DIST of the University of Genoa, and IWT (Intelligent Web Teacher), developed at the DIIMA of the University of Salerno, each of them characterized by the use of ontologies and semantic technologies in order to support instructional design and personalized learning processes. The former aims to develop a learning resources and instructional paths authoring tool based on a logical and abstract annotation model, created with the goal of guaranteeing the flexibility and personalization of instructional design, the reusability of teaching materials and of the related whole knowledge structures. The latter represents an innovative e-learning solution able to support teachers and instructional designers to model educational domains knowledge, users’ competences and preferences by a semantic approach in order to create personalized and contextualized learning activities and to
allow users to communicate, to cooperate, to dynamically create new content to deliver and information to share as well as enabling platform for e-learning 2.0.

1 Introduction

The evolution of Semantic Web (SW) technologies during the recent years has made ontology-based systems an interesting framework for supporting learning processes (http://www.w3.org). Despite this, according to Mizoguchi and Bourdeau (2000) some open issues need to be addressed, such as a deep conceptual gap between authoring systems and authors; authoring tools are neither intelligent nor particularly user-friendly; knowledge and components embedded into Intelligent Web-based Education Systems (IWBES) are rarely sharable or reusable; there is a gap between instructional planning for domain knowledge organization and tutoring strategy for dynamic adaptation of the IWBES behaviour.

Moreover, a relevant problem concerns the fact that there is no canonical representation of knowledge structures, because any of them can be seen from different points of view, showing different structures. As Ohlsson (1987) highlighted, this fact has such relevant implications for authoring systems that it should be stated as the

“Principle of Non-Equifinality of Learning”, according to which “the state of knowing the subject matter does not correspond to a single well-defined cognitive state. The target knowledge can always be represented in different ways, from different perspectives; hence, the process of acquiring the subject matter have many different, equally valid, end states”.

It should be clear that a large number of possible instructional paths representations exists and the problem is to find the appropriate content and presentation for the specific learners’ profile and the specific educational purpose which have been defined (Baker, 2000).

Another issue that needs to be addressed is related to knowledge modeling by means of metadata definition standards. These latter 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 (Stojanovic et al., 2001).

Several systems have been developed to handle learning resources by means of SW technologies, they are usually named SWBE (Semantic Web Based Education) Systems. Their components are generally related mainly to the roles of users (teachers, learners, authors, groups, developers), to educational resources,
to the environment, to its interface and to the functionalities it does offer. As previously said, the Semantic Web appears to offer innumerable advantages as regards to “traditional” Web and this is the reason why SWBE Systems created great expectations and a rapid deep-change was expected with the development of the so-called Educational Semantic Web (Anderson & Whitelock, 2004). Actually, this change has not taken place due to the lack of availability of shared ontologies, ontology-based semantic annotation of learning resources, and educational services based on semantic technologies. Despite this, the first interesting results are now available in contexts characterized by the application of Artificial Intelligence methodologies and approaches into WBE Systems (the so-called IWBE Systems), which have as main components: ontologies, pedagogical and tutoring agents, semantic-based tools and services (Devedzic, 2003).

In this paper, some issues related to IWBEs are discussed in the introduction to two research projects. The former, CADDIE (Content Automated Design & Development Integrated Editor), was developed at the ELKM – E-Learning & Knowledge Management Lab. – DIST, University of Genoa (www.elkm.unige.it). The latter, IWT (Intelligent Web Teacher) was developed at the CRMPA – Centre of Research in Pure and Applied Mathematics – DIIMA, University of Salerno (www.crmpa.it).

CADDIE is an instructional resources and paths designing tool based on a knowledge structure representation model, named ECM - Educational Concept Map (Adorni et al., 2009). IWT allows to model educational domains knowledge, users’ competences and preferences by a SW approach in order to create personalized and contextualized learning activities, and allows users to communicate, to cooperate, to dynamically create new content to deliver and information to share as well as enabling a platform for e-learning 2.0.

2 CADDIE (Content Automated Design & Development Integrated Editor)

The goal of the CADDIE project is to realize a complete framework for the macro and micro design of lessons, educational paths and educational aids. The theoretical background behind this model comes from the studies on instructional content design (Merril, 1999; Gagné & Briggs, 1990), knowledge representation and modeling of learning experiences with an important influence of the Educational Modeling Language (Rawlings et al., 2002), the semantic web (Berners-Lee, 2001; Shadbolt et al., 2006) and the subject matter framework (Adorni et al., 1981).

The CADDIE project, grown out of these studies, started with the design of a model for learner profiling and a logical and abstract annotation model
named Educational Concept Map (ECM).

Target group analysis is the first step in order to design an effective set of learning materials. This activity allows you to get useful information to design a “personalized” or “individualized” kind of learning. We underlined the previous two words because they are not synonyms: Baldacci (2005) claims a different meaning for the two approaches: with “personalization” we refer to didactic actions aimed at adding value to the individual in relation to the group, while with “individualization” we refer to individual learning activities where the student works alone on his/her own learning materials developed ad hoc according to his/her learning attitudes and styles. Notwithstanding this distinction, in this paper the two words are used as synonyms with a certain stress on “individualization”.

The main international standards for target analysis are the IEEE LTSC Personal Private Information Standard (LTSC-IEEE, 2001) and IMS Learner Information Package (IMS, 2006). These standards indicate a number of items aimed to describe all the student’s characteristics: personal, cognitive, motivational and his/her competencies. Cognitive characteristics are the most relevant in order to choose the right didactic strategy and to design the most suitable learning materials (Adorni et al., 2008).

Two are the main dimensions of the cognitive characteristics (Crispiani, 2004): “kind of intelligence” and “learning style”. In order to define the concept of “kind of intelligence” we refer to the Gardner’s well-known theory of multiple intelligences. Gardner (1983) identifies the following: Linguistic Intelligence, Logical-Mathematical Intelligence, Spatial Intelligence, Musical Intelligence, Bodily-Kinesthetic Intelligence, Interpersonal Intelligence, Intrapersonal Intelligence, Naturalistic Intelligence.

Less partitioned results the “learning styles” classification (Crispiani, 2004) that is formed by the following three elements: Analytic Learning style, Interactive Learning style, Introspective Learning style. McKenzie (2006) found a correspondence between learning styles and kind of intelligences. He proposes the following combinations:

- Analytic Learning style with Logical-Mathematical Intelligence, Musical Intelligence, Naturalistic Intelligence;
- Interactive Learning style with Linguistic Intelligence, Bodily-Kinesthetic Intelligence, Interpersonal Intelligence;
- Introspective Learning style with Spatial Intelligence, Intrapersonal Intelligence.

The target analysis model proposed for CADDIE aims at a content design activity conform to the ECM model: the identification of the learning style allows to choose the most suitable didactic strategy and the awareness
of the student’s kind of intelligence allows the designer to choose and design both the optimal didactic techniques and the most effective learning contents. (Battigelli & Sugliano, 2009). Just an example: for a student with an Analytic Learning style and a Logical-Mathematical Intelligence, the model suggests a learning technique based on tutorials and learning materials that favor graphic representations.

In order to assess the target group analysis we propose to use the test built by McKenzie (2006). This test allows to use a unique analysis tool in order to obtain information about the user cognitive attitude, with reference to the kind of intelligence and the learning style.

The ECM is a logical and abstract annotation model created with the aim of guaranteeing the reusability not only of teaching materials, but also of knowledge structures (moving the generalization level from the contents to the definition of the contents’ schema). Once defined the learner is profiled, the model suggests identifying, within the discipline’s subject matter, the key concepts and their relationships so as to identify the most effective strategies of contents presentation and to support the activation of meaningful learning processes.

The educational objectives, according to the model, can be represented as SingleObjective if they are not decomposable into sub-objectives (single objectives) or ComposedObjective, if they are composed by two or more sub-objectives (composed-objectives). Single units of learning including topics (SingleUoL) will be associated to the first class; while composed units of learning (ComposedUoL) including SingleUoL or ComposedUoL will be linked to the second class. The model is based on a hierarchical and recursive organization (through the is-a relationship) of learning objectives to which corresponds a layered structure (n levels with n integer positive) of contents as we can see in figure 1.

Fig. 1 - Educational Concept Map schema
Moreover, it is worth to point out that the relationship between an objective and a unit of learning is always necessarily a 1:1 association (one SingleUoL corresponds to one SingleObjective). An Educational Concept Map can be therefore defined by the syntax:

\[
<ECM> ::= <UoL> + \\
<UoL> ::= <SingleUoL> | <ComposedUoL> \\
<SingleUoL> ::= <Topic> \\
<ComposedUoL> ::= <ECM> \\
<Topic> ::= PrimaryTopic | SecondaryTopic
\]

The BNF grammar shows the composite structure of the ECM model, but it doesn’t define any pedagogical relationships between Objectives (as for the relationships between the Unit of Learning). These relationships have the form of:

\[
R(UoL_1, UoL_2)
\]

In that way we obtain the logical and chronological scheme of educational objectives. It is possible to identify and organize the schema of contents based on a taxonomic learning units organization (in fact, each objective is pursued through the corresponding UoL). Then, for each unit of learning, the ontology needs to specify the topics, the key-concepts on which the UoL is focused. Topics can belong to the two following classes (Adorni et al., 2009):

- **PrimaryTopic**: this class identifies the “prerequisites”, in other words the concepts that the student must know before attending the course (the set of these topics, which will not have instructional resources associated, represent the knowledge-requirement of the course);
- **SecondaryTopic**: this class identifies the concepts which will be explained in the course of the unit of learning (this kind of topics will have specific learning materials associated).

The Rel relationship which establishes a connection between a primary topic and a secondary topic is named *is-primary-topic-of*, while the relationships between *SecondaryTopic* are:

- **is-requirement-of**: it identifies a transitive and propaedeutic association among two or more topics (e.g., it may be used with the aim of specifying the logical order of contents);
- **is-related-to**: it identifies a symmetric association among closely related topics (e.g., it may be used with the aim of creating learning paths without precedence constraints);
- **is-not-related-to**: it identifies a symmetric relation of indifference between two or more topics (e.g., it may be used with the aim of making explicit the absence of association among topics);
- **is-suggested-link-of**: it identifies not-closely related concepts (e.g. this relationship type may be used in order to suggest in-depth resources, internal or external to the contents repository).

These relation types have been selected with the aim of allowing teachers and instructional designers to create different learning paths (with or without precedence constraints among topics).

More formally, given a set \( \{t_1, \ldots, t_n\} \) of topics \( T \), we can define the subset of prerequisite \( P (P \subset T) \) as:

\[ t_j \in P \leftrightarrow \exists t_i \in: \text{is-primary-topic-of}(t_j, t_i) \text{ with } i = 1, \ldots, j-1, j+1, \ldots, n. \]

Similarly, it is possible to define the “learning outcomes” \( L \) subset of \( T (L \subset T) \) as follows:

\[ t_k \in L \leftrightarrow \exists \neg t_i \in: \text{is-requirement-of}(t_k, t_i) \text{ with } i = 1, \ldots, k-1, k+1, \ldots, n. \]

1. Subsequently to the ECM design, it will be possible to associate educational resources to the single nodes (for example, Learning Objects, text documents, audio and/or video files, etc.). CADDIE handles the ECM using the XML Topic Maps (XTM) standard (Adorni et al., 2007), that can be used for web browsing based on topic connections or it can generate a sequence of topic for printable documents or traditional lesson planning.

2. In this latter case a Suggested Paths Strategy is necessary, by means of is-requirement-of relationships: for every path between two nodes, and for every node of the above path, reading the string from left to right do:

3. insert before a node on the string all the possible sub paths from PrimaryTopics to that node, reading the graph from top to bottom;

4. prune: reading the new string (as generated at step 1.a) from left to right, delete possible node repetitions (keep only the first letter of a node);

5. compute the real distance \( \Delta S_{ij} \) on the string of every contiguous node (nodes with \( \Delta = 0 \)) of the graph;

6. insert suggestions: if a string exists with \( \Delta S_{ij} > \Delta_{\text{max}} \) insert a suggestion for a Topic Aider of the node I before the node J;

With this algorithm we obtain all the suggested paths that are ordered following the criteria of: number of \( \Delta S_{ij} > 0 \), and on equal terms for increasing \( \Sigma \Delta S_{ij} \). The choice to have not a single path (maybe the “best” according to some minimum principle) but a list of paths to suggest to the author leaving the final choice to the author him/herself, is to answer to the non-equifinality problem posed by Ohlsson (1985). The “suggested” order lists is on the base
of the principle of reducing as much as possible the distance of two topics of the list which are contiguous on the graph.

Nowadays CADDIE is a prototype realized under the open source philosophy.

3 IWT (Intelligent Web Teacher)

Intelligent agents of IWT operate on three main modules: the didactic knowledge, the student model, some planning procedures. (Albano et al., 2006; Capuano et al., 2008).

The educational knowledge is represented through different abstraction levels. The lowest level is composed by Learning Objects. Learning Objects must be indexed in order to let the engine know what each one of them is about and how they can be used during the learning process. This is done by a second abstraction representation level (Metadata). A Metadata is a collection of attributes about a Learning Object (LO) describing some features such as its type (text, simulation, slide, questionnaire,...), the required educational level (high school, university,...), the language, the interactivity level and so on. Finally, a third abstraction level (called Ontology) is used to represent educational Domain Concepts and their relations. A Domain Concept (DC) is a concept belonging to the described educational domain and can be possibly explained by one or more LOs. Typical relations among concepts are: Has Part, Is Required By, Suggested Order, to indicate, respectively, a hierarchical relationship and a hard or light constraint on the learning order of two concepts.

The ontology describes semantics under the content by graphically repre-
senting concepts and relationships among them (Gaeta et al., 2009). In fact some concepts are required by another one when this one has to be included in a course. A course has principally a learning path extracted from the ontology i.e. a sequence of concepts from the domain dictionary.

![Diagram](image)

**Fig. 3 - The learning path is a sequence of educational domain concepts**

When the system has to deliver a course, its learning path has to be translated in a presentation. Each concept of the learning path will be covered by at least a learning object in the presentation.

The student model collects information about the student’s Cognitive State and Learning Preferences. The planning procedures are able to automatically create a course satisfying all the student’s learning requirements taking into account both her/him cognitive state and learning preferences.

Given a set of Learning Objectives chosen by the teacher on the educational domain ontology, IWT is able to generate the best Learning Path for each student starting from his Student Model. Different students with the same Learning Objective will so have different courses generated by the system.

IWT supports the integration of resources, tools and services (Capuano et al., 2009). Thanks to this feature, Web 2.0 and, then, e-Learning 2.0 main aspects and tools became IWT extensions early. IWT could be seen as a complete e-Learning 2.0 solution because it supports:

- Learner centred approach: IWT foresees the learner at the centre of the learning process;
- Personalization and Contextualization of the learning experiences: IWT foresees the importance of those aspects; its own model, process and services help to personalize learning experiences;
- The importance of educational theories in e-Learning;
The importance of semantics and knowledge in learning: semantics is horizontal to all the IWT services and annotation of resources is central.

The adoption of a service oriented view. Here the IWT model is, of course, based on the concept of Service Oriented Architectures (SOA). IWT integrates, of course, LMS’s and LCMS’s functionalities and a wide set of Web 2.0 tools:

- e-Portfolio. The portfolio stores, in an organized manner, personal information, learning style, cognitive state, tracks of learning activities in which a student is involved or has been involved, etc. Owners of each e-portfolio will choose which data should be private and which should be public.
- Blogs. In IWT students can share their ideas about fixed or open topics. Educators can fix and make explicit their knowledge about specific arguments.
- Podcasts. A simple way to capture and spread video/audio learning content. Podcasts can be used to record and disseminate teachers’ lectures. Podcasts can also be used as deliverables of students’ tasks.
- Wikis. A widespread mechanism used to construct structured knowledge collaborating with other people. In IWT Wikis could be used by groups of learners in order, for instance, to collaboratively construct ontologies. The produced artefacts could be evaluated by a teacher to assess learners’ tasks.
- Social Networking and Bookmarking. Users may keep in contact with each other in an informal way; they may set up a study group formed by learners with the same learning goals; they may use this service to find people having same skills, preferences, learning styles, interests, etc.
- Knowledge Forums. Where people may post questions and answer and tag them by semantic indexing and rate them by informal-intentional mechanisms.
- Shared Areas. Where people may share content, download others’ content, tag it, post rating comments and feedbacks.
- RSS Feeds. By means of them, people may publish and collect quickly and easily information on state, activities, interests, etc.

All user generated content and Web 2.0 services (O’Reilly, 2005) complete the educational offer of IWT. They are semantically annotated by employing available semantic technologies and architectures as well as other resources. It allows IWT to select learning objects and service to create the best learning activities for users getting them to their own educational goals.

IWT has been employed in many contexts. In enterprises, universities and schools (more than 50.000 users) IWT has been experimented to demonstrate
the benefits of personalization based on Semantic Web approach in on-line learning and the advantages of the Web 2.0 tools availability.

Personalization makes it possible for users to feel better. Courses eliminate real lacks and usually go on till the learning objectives have been reached by the users. Personalization selects also the best resources for them, IWT is able to match preferences and didactical approaches.

Web 2.0 is able to involve learners in their own learning activities, to create enhanced social networks to communicate and collaborate with other users facing the same problems, subjects and courses, to share content, to receive feedbacks and comments, to be in contact by means of videoconference tools, be connected through mobile devices.

Such experimentation (Capuano et al., 2008) involved groups in universities and enterprises. We divided each group in two subgroups, we allowed everybody to use technologies for a period and we compared the results at the end.

As you can see from the graph the personalization of IWT allows users to learn in a more effective way. The competences gained from users have been divided in three parts: high, medium and low level. High level competences without personalization (without LIA, the bar in the middle of Fig. 4) remain the same, meanwhile they grow with personalization (with LIA, the bar on the right of Fig. 4). Medium level competences without personalization (without

![Fig. 4 - Experimentation results on personalization: on the left competences before learning activities, in the middle competences after e-learning activities in IWT without personalized courses (LIA is the intelligent agent able to do them), on the right competences after e-learning activities in IWT with personalized courses.](image)
LIA, the bar in the middle of Fig. 4) grow enough, meanwhile they grow more with personalization (with LIA, the bar on the right of Fig. 4). Low level competences without personalization (without LIA, the bar in the middle of Fig. 4) decrease enough, meanwhile they decrease closely to zero with personalization (with LIA, the bar on the right of Fig. 4).

Conclusions

In this paper, two different research projects have been introduced: CADDIE (Content Automated Design & Development Integrated Editor) developed at the DIST, University of Genoa; and IWT (Intelligent Web Teacher) developed at the DIIMA, University of Salerno; they were designed in the context of educational systems in order to accomplish the AAAL: “Anytime, Anywhere, Anybody Learning”. These projects have been discussed here together due to the fact that both of them, even adopting different methodologies, refer to Intelligent Web-based Education Systems and represent two good practices in the context of the application of ontologies and semantic technologies to support instructional design and personalized learning processes.

At present, the researchers involved in these projects are exploring different research hypotheses with the aim of developing their systems further. With reference to CADDIE, the line of research is addressed to integrate it into a planning system able to develop, once prerequisites and educational goals are defined, instructional paths and related resources by means of an automatic retrieval process of Units of learning available on the Web.

The future evolution of IWT will foresee the convergence of formal and informal approaches for e-learning (following the trend of collaborative and social applications) by exploiting the Personal Learning Environment as an integration paradigm and the competency-driven learning also by focusing on the competency-based management in the enterprise context. In fact the problem to face is not to employ Web 2.0 tools by focusing only on new technologies, but to collocate them in instructional design aiming, thus, to lead processes, to arrange activities, to avoid the “Phoenix effect” of e-learning 1.0 and the learning chaos where technology-based empowerment could eventually carry on. The next future Semantic Web approach in IWT will improve the way to index content, to share knowledge, to create courses and, finally, to do e-learning by finding a convergence among Semantic (ontologies, taxonomies, controlled vocabularies) and Social (tagging activities, folksonomies, rating, reputation) aspects.
REFERENCES


Baldacci M. (2005), *Personalizzazione o Individualizzazione?*, Trento: Centro Studi Erickson.


IMS Learner Information Package (LIP) - http://www.imsproject.org/profiles/ (verificato il 26/03/2010).