

An Ontology-Based Approach for Sharing Digital Resources in Teacher Education

Serena ALVINO^a, Stefania BOCCONI^a, Pavel BOYTCHEV^{b,1}, Jeffrey EARP^a, Luigi SARTI^a

^a*Istituto per le Tecnologie Didattiche – C.N.R., Genova, Italy*

^b*NIS-SU, Sofia University, Sofia, Bulgaria*

Abstract. Teacher Education (TE) is a dynamic, lifelong process that needs to fully embrace innovation and assume a broader European perspective. The EC-funded Share.TEC project aims to provide enhanced, culturally-aware access to TE-related resources across Europe by means of a federated resource brokerage system whose semantic core is the proposed Teacher Education Ontology (TEO). This paper describes the rationale for an ontology-driven approach, gives an overview of TEO's multi-layered structure for addressing multicultural and multilinguistic issues, and presents some aspects of the TEO implementation that allow for language-independent conceptualization and multidimensional hierarchal searching and filtering. Other TEO features are also discussed, including the support for dynamically generated user interface and system stability against ontology modifications.

Keywords. teacher education, ontology, ontology-based Information System, multicultural semantics

Introduction

Teacher Education (TE) is a lifelong learning process that is central to Lisbon Strategy efforts towards the building of a European knowledge society. However, the TE field has been generally slow to embrace innovation, much less to generate it, and has yet to assume a broad European perspective. Major hurdles stand in the way: TE practice is usually geared to meet the specific requirements of national systems that are linguistically and culturally bound; TE communities (even virtual ones) tend to focus on the immediate locus; hesitancy persists in embracing digital culture, with only patchy adoption of ICT and scarce sharing of digital resources.

Providing impetus for innovation within initial and in-service TE is the goal of the EC-supported Share.TEC² project. Share.TEC has undertaken to build an advanced user-focused system dedicated specifically to fostering a stronger digital culture in the TE field. This system is to aggregate metadata describing TE-related digital resources located Europe-wide; provide personalized, culturally-sensitive brokerage for the retrieval of relevant digital content; support the development of a Europe-wide perspective among those working in and with the TE community.

¹ Corresponding Author: Pavel Boytchev, KIT, Faculty of Mathematics and Informatics, 5 James Bouchier Blvd, Sofia, 1164, Bulgaria; E-mail: boytchev@fmi.uni-sofia.bg.

² Share.TEC - SHARing Digital RESources in the Teaching Education Community, eContentplus programme (ECP 2007 EDU 427015); <http://www.sharetecproject.eu/>.

In pursuit of these objectives, an ontology-driven approach has been adopted. The semantic core at the heart of the proposed Share.TEC system is a Teacher Education Ontology (TEO), which is currently being developed by partners in the Share.TEC project in collaboration with international experts. The scope of this ontology has been set on concepts relevant to the domain of Teacher Education, with particular regard for aspects considered pertinent to the sharing of digital resources and practice among potential members of the Share.TEC community, namely teacher educators, teachers, academic/educational publishers and content developers.

The purpose of TEO within the Share.TEC system is to provide:

- pedagogical characterization of digital content;
- representation of user profiles and competencies as a backbone for cultivating TE communities;
- a basis for multilingual and multicultural functionality;
- support for personalized interaction with adaptive user applications;
- support for the implementation of recommending functions.

There are a number of reasons why an ontology-based approach has been adopted for the Share.TEC system and platform. The chief among these is to permit the sharing of concepts among people. Share.TEC's European perspective necessarily means that its users will bring to the community different languages and cultures. What's more, the TE field includes people with very different backgrounds, ideas and assumptions. In such a situation, effective communication and shared understanding can be difficult to achieve. Accordingly, TEO seeks to reduce conceptual and terminological confusion by identifying and properly defining a set of concepts (and their relations) relevant to TE in Europe. The result should be a non-ambiguous and consistent vocabulary for identifying those concepts, and a framework on which culturally and linguistically diverse versions of that vocabulary can be mapped.

As an integral part of the Share.TEC system, TEO will support adaptive user interfaces, and will inform services that use reasoning techniques. This should lead to the implementation of inferential search engines, advanced ranking solutions and flexible representation of user profiles. Importantly, TEO has also provided the basis for the definition of a common metadata model for describing TE-relevant digital resources.

1. A Domain Ontology of Teacher Education

1.1. TEO Definition and Development

Domain ontologies, upon which Information Systems are subsequently based, can be considered as repositories of knowledge that allow accumulation and systematization of knowledge. According to Allard et al. [1], "*Domain ontologies should also be conceived as use-neutral, in the sense that they are meant to serve as a foundation. Building on this foundation, different problems can be tackled, various applications derived, knowledge bases built. Consequently, domain ontologies should be relatively stable and aim to be a long-lasting conceptual structure*".

As previously mentioned, the present work on domain ontology for Teacher Education seeks to capture those concepts of the TE world that are relevant for sharing digital resources among practitioners. It is also to provide a framework for mapping multicultural and multilinguistic semantics.

Figure 1 describes the process of TEO development.

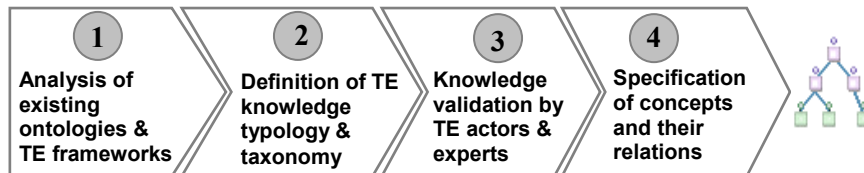


Figure 1. Phases in TEO's development process

As shown in Figure 1, TEO is grounded in existing research, especially that of Mizoguchi *et al.* [1], [2] and Guarino [3]. Specifically, it draws on three main models of reference: the OMNIBUS ontology³, whose domain is education; the LORNET competency modelling ontology⁴; and the POEM learning object content model [4]. Other relevant sources that have influenced TEO's design include DOLCE⁵, ONTOURAL [5], ALOCOM⁶, PROTON⁷ and user modeling ontologies [6], [7].

1.2. Basic Concepts of TEO

In its current version, TEO defines **162 classes** and **78 properties**. Figure 2 presents the main concepts identified.

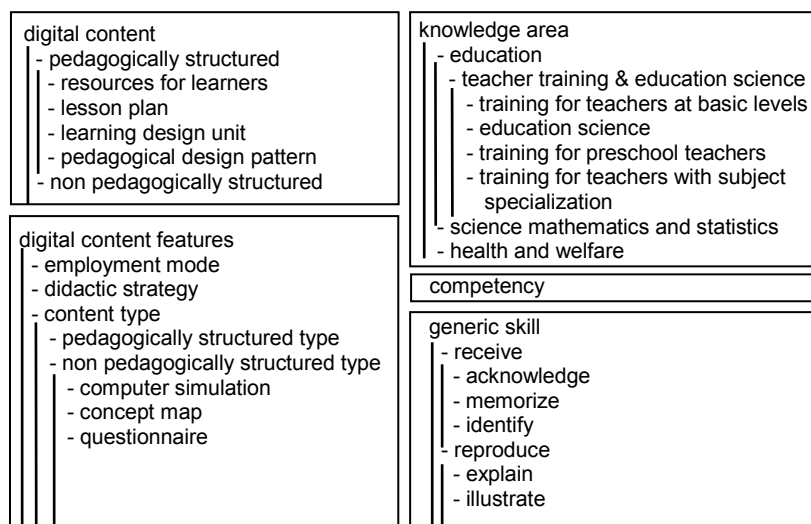


Figure 2. Overview of TEO's main concepts

³ <http://edont.tee.jp/omnibus/doku.php>

⁴ <http://www.lornet.org>

⁵ <http://www.loa-cnr.it/DOLCE.html>

⁶ <http://ariadne.cs.kuleuven.be/alocom/>

⁷ <http://proton.semanticweb.org/>

Digital content refers to educational resources and artifacts closely related to the concept of the “learning object” [8]. Depending on its nature, a *digital content* artifact can be categorized as (i.e. “is-a”) *pedagogically structured* or *non-pedagogically structured type*. *Resources for learners, lesson plans, learning design units* and *pedagogical design patterns* belong to the *pedagogically structured* category.

The *digital content* concept is also defined by other characteristics including *employment mode, didactic strategy* and *content type*. These act as *digital content features*, i.e. they are related to *digital content* instances via a *part of* relationship. A *computer simulation, a concept map* and a *questionnaire* are all examples of *non-pedagogically structured content type*.

Knowledge areas consist of topics drawn from the EUROSTAT [9] taxonomy of education and training. This classification was adopted as a reference due to its European perspective and pertinence to the TE domain⁸. The knowledge area hierarchy allows digital content to be described in terms of discipline and permits specification of the user’s areas of interest. Association of a knowledge area and a *generic skill* generates a *competency* concept.

According to Paquette [10], competencies

“[...] *link skills and attitudes to knowledge required from a group of persons and, more generally, from resources*”.

In TEO, competencies are related with the *digital content* concept: this allows resources to be described and classified according to the specific competencies they address.

Following Paquette’s approach to competency modeling [10], *generic skills* represent basic cognitive processes expressed as simple action verbs like *receive, reproduce, perceive, analyze, synthesize*. These concepts are arranged in a hierarchy: for instance, *acknowledge, memorize, identify* are lower levels of the *receive* action. At the top level, this hierarchy spans a continuum of “cognitive complexity”.

Finally, we also considered the *role* concept, which draws on Mizoguchi’s model [2].

Once basic concepts were defined in TEO, a knowledge validation process was carried out by TE actors & experts (see step 3 in **Figure 1**) to progressively improve TEO’s conceptual framework and its technical implementation. Some important requirements emerged from development of Share.Tec’s technical integration study and system architecture specification, so a new release of TEO has been produced to fulfill these needs.

2. TEO Implementation

TEO is designed to represent concepts and build relations between entities defining the domain of Teacher Education. This in itself is a positive step towards defining a consistent and complete picture of this domain, but TEO also contains important information that can be used by the Share.TEC software application.

The main contents of the Share.TEC repository are data, harvested from external repositories or provided by Share.TEC community members that are related to digital

⁸ For interoperability purposes EUROSTAT has been mapped against Dewey’s Decimal Classification: <http://www.oclc.org/dewey/>.

content. Along with these primary data, the main repository contains an online dynamic representation of TEO that supports core system services and features like:

- language-neutral concept-oriented data
- hierarchal searching and filtering
- dynamic multilingual user interface
- stability and system independence with respect to future changes in TEO.

2.1. Internal Logical Structure of TEO Representation

The internal logical structure of a TEO entity is designed with a minimalistic approach in mind – the simplest structure that facilitates all required functionality. Each TEO entity is represented as an individual node that is interconnected with other nodes through relations and that contains a list of translations of the concept represented.

The internal physical structure is more complex and is not discussed in this paper. It defines a wider spectrum of relations between ontology entities and contains additional data in order to allow (a) complete reconstruction of the ontology into a valid OWL file; and (b) support for extended functionalities like reasoning.

The structure of the “Medicine” node is shown in **Figure 3**. The node represents a language-neutral concept. It contains its verbal representation in three languages and is connected with other nodes through parent-child relations. This structure is sufficient to support conceptualization, hierarchies, multilingualism and ontology stability.

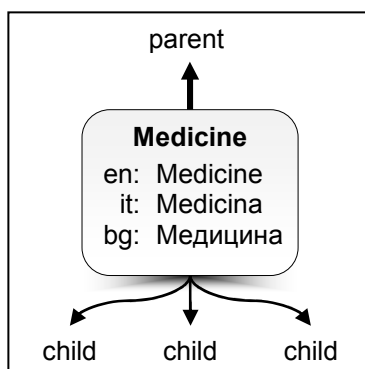


Figure 3. Internal structure of a TEO node

While this logical representation does not distinguish between classes, subclasses and instances, the distinction is made in the physical representation. TEO hierarchies are constructed by defining parent-child relationships between various nodes. This consolidates all TEO nodes into a single tree-like data structure that represents the domain knowledge of Teacher Education. Additional cross-branch relations are also represented, but these are not discussed here. **Figure 4** shows a vertical slice describing the complete path from the top concept, TEO, down to Cardiology. The path goes through Knowledge Area node, which is the root of the Knowledge Area hierarchy within TEO.

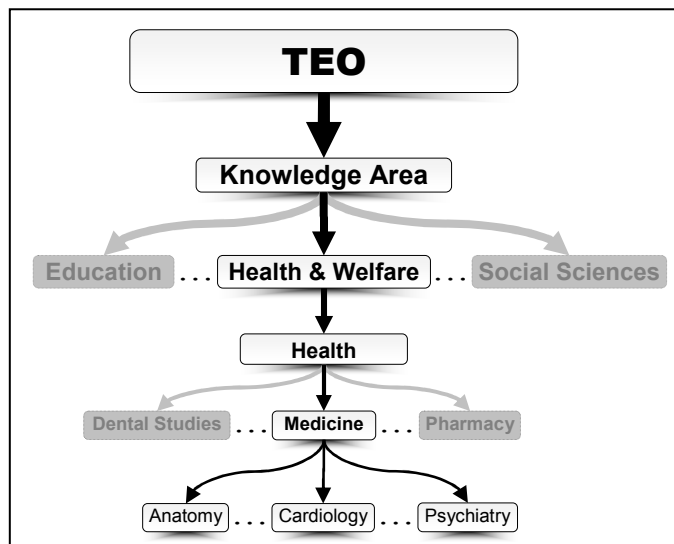


Figure 4. A vertical slice of TEO representing the full path down to Cardiology

2.2. Multilingual Support

Every node contains a set of translations of the node's concept into system-supported languages. This information is used when processing multilingual data from repositories across Europe. Whenever incoming data contains a concept expressed in a native language, Share.TEC scans TEO to find the corresponding node. When such data are processed, their texts are replaced by references to conceptual nodes. This makes the internal representation of data language-independent and links various translations of the same concept – **Figure 5**.

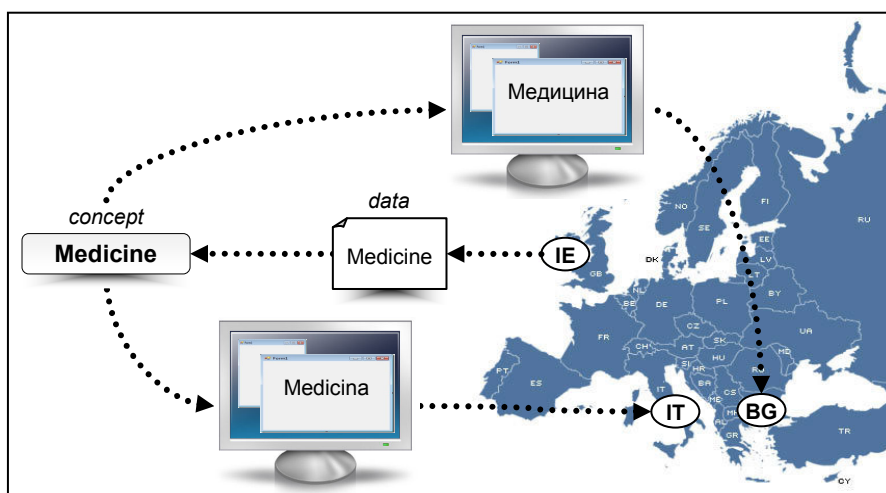


Figure 5. Native languages and language-neutral conceptualization

The same translations are used in the opposite direction, i.e. to translate concepts into users' native languages. The end-user system is multilingual and texts are translated on-the-fly while the screen forms are dynamically built. The original data harvested from, say, an Irish repository may contain the word "Medicine" but the same data viewed by Italian or Bulgarian users will be displayed as "Medicina" or "Медицина".

The current node structure can handle synonyms by providing more than one translation of the concept in a given language. Additionally, each node may have a description, which is available in all supported languages.

Translations are actually used not only for importing and displaying data, but in other activities, like searching and filtering. For example, a Bulgarian user may define a search criterion in Bulgarian (e.g. Knowledge Area="Медицина"); the system will match this to the node representing the Medicine concept and will find all data referring to that concept, irrespective of language.

2.3. Multidimensional Hierarchal Data Space

The hierarchal structure of TEO contributes to searching and filtering capabilities by expanding functionality. Instead of searching for a specific low-level concept like Cardiology, the user can set a higher level criterion such as Medicine. The result will be that all data referring to any of the 28 concrete concept instances under Medicine will match that criterion – from Anaesthesiology via Neurology and up to Psychiatry.

If the parent of Medicine (i.e. Health) is selected, then all medicines, dental studies, medical diagnostics, treatment technologies, nursing, caring, pharmacy, therapy and rehabilitations are selected.

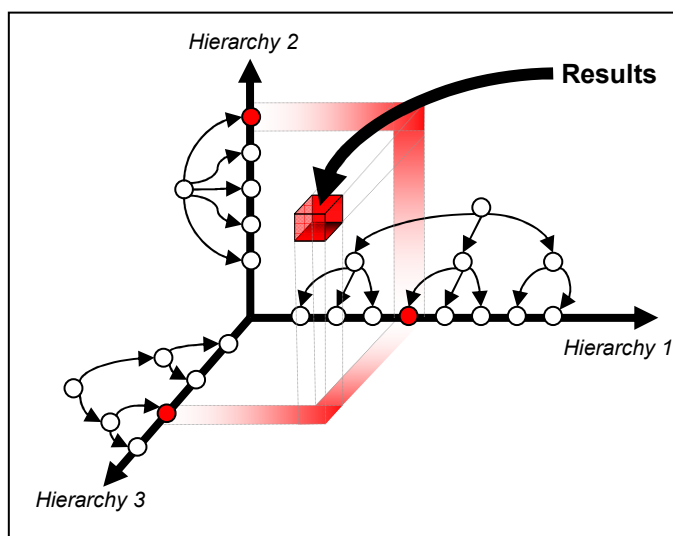


Figure 6. Multidimensional data searching

TEO hierarchies can be represented as axes in a multidimensional data space. Note that these axes are not mathematically continuous but are discrete sequences of concepts. **Figure 6** shows an example of a 3D data space defined by three hierarchies.

Solid circles represent selected entities. If all data are positioned in this 3D space, then applying filtering criteria will cut out a virtual box from all data.

All data elements inside this box are the user's search or filtering results. The multidimensional interpretation of Share.TEC data enriches the way users perceive TEO. They can "slide" along TEO axis by axis, slicing the data space in their preferred way.

Users are free to select any number of TEO axes and in any order (not just three as in **Figure 6**). Moreover, it is possible to select any non-leaf node from any hierarchy. This feature enables the user to limit the scope of the result not only to individual elements, but also to groups of related entities.

TEO has already captured and classified relationships between concepts. This information is used to define the boundaries of each group of related entities at any level of classification. Groups of related entities can be retrieved simply by accessing an upper-level node from the corresponding hierarchy – **Figure 7**.

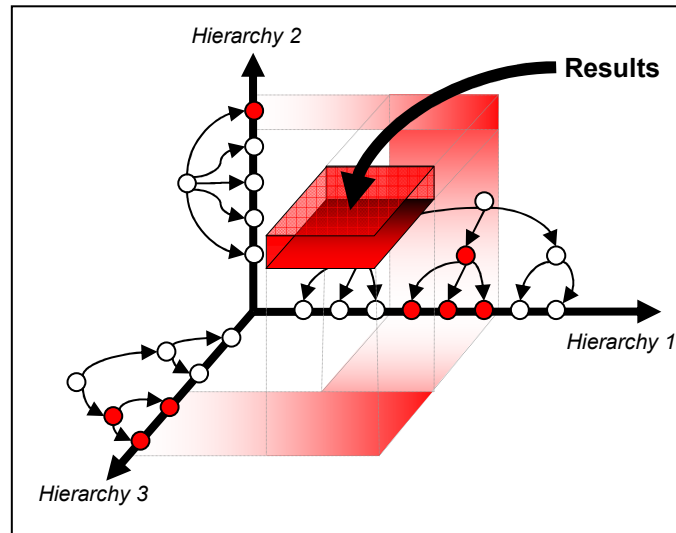


Figure 7. Hierarchical searching of data bound by a virtual searching box

The effect of this multidimensional and hierarchical approach is that the user can broaden or narrow the filtering criteria in a natural and domain-aware way. Axes corresponding to parameters which are not of interest to the user are totally ignored. For example if the user does not want to impose a filter on language, i.e. any language is acceptable, then instead of selecting the root node of all languages it is much easier just to ignore the language axis.

The multidimensional approach does not only utilize the taxonomies in TEO. A class together with its properties and the properties thereof (and so on) can also be treated as a hierarchy and thus be subject to multidimensional search and navigation.

2.4. System Sensitivity and TEO Modifications

An important element of the design of any data-driven system is to analyze its sensitivity with respect to modifications in the data. Our proposal is designed to be

non-sensitive or stable with respect to two major classes of modifications (stability meaning that if data are changed, then there is no need to modify the software).

The first class of modifications is adding a new language. Language-specific data is stored inside TEO nodes, so a new language will only add new properties and will not change the structure of TEO. The software application retrieves language information dynamically, so whenever a new language is added, it can be used right away. This means that while a new language is being added to the Share.TEC system, services can continue without interruption.

The second class of modifications affects the structure of TEO. If new instances or subclasses are added to any of the hierarchies, the system will use them right away, without any need for upgrading. However, if the modification goes beyond the boundaries of existing hierarchies, i.e. a new hierarchy is added, then the system will need to undergo major changes, because new data fields should be imported and/or analyzed.

3. Future Work

This paper has presented a domain-ontology for Teacher Education (TEO). Some of the major ideas underpinning the adopted approach were discussed, as were technical aspects regarding the implementation of TEO and related services.

Although minimalistic, the proposed implementation of TEO covers a wide range of features required for appropriate functioning of Share.TEC portal. It has been found that this implementation approach might be suitable for representing multicultural diversity in TEO. This is especially important where two or more cultures have their own specific views of TEO which are not compatible.

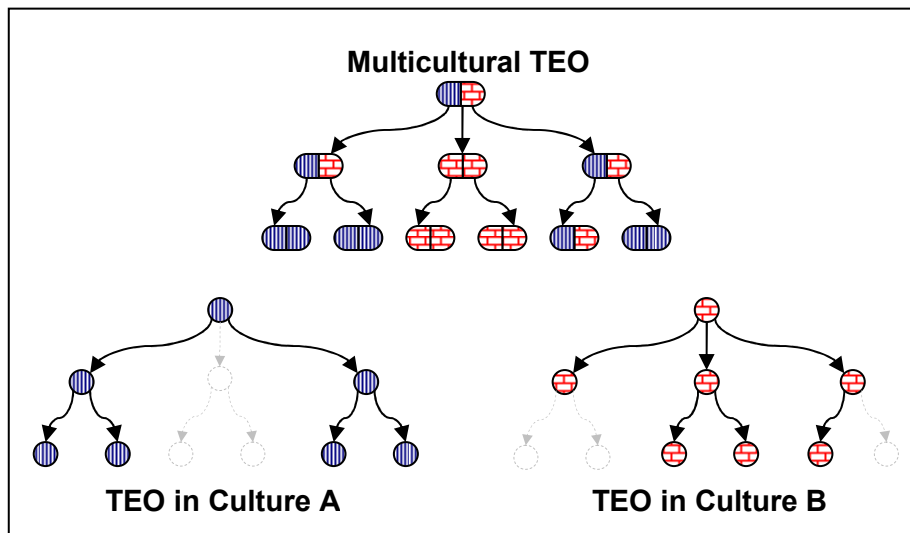


Figure 8. Multicultural coverage

Multicultural support goes beyond multilingual support, as it deals with those elements for which translation is problematic, i.e. it cannot be resolved with direct correspondence. The basic assumption is that when a node contains at least one

translation in a given language, then the corresponding concept exists in the culture based on that language. Thus the same data space will be seen in different ways by different users – **Figure 8**.

Although the projection of TEO onto a given culture may hide some of the concepts (especially those which do not exist in that culture), users are still able to view and work with the complete data space if they remove the language/culture filter. However, in this case they risk encountering unfamiliar and untranslatable concepts.

Acknowledgements

All the authors contributed on an equal basis to the conceptual design of this paper. ITD authors were mainly responsible for the introduction and first section, while Pavel Boytchev of NIS-SU was mainly responsible for the second and third sections.

References

- [1] D. Allard, J. Bourdeau, R. Mizoguchi, Towards modeling knowledge of cultural differences and cross-linguistic influence in Computer-Assisted Language Learning (CALL). In: E. Blanchard & D. Allard, *Proceedings of the Culturally Aware Tutoring Systems (CATS) Workshop*, ITS Conference 2008, Montreal, Canada, June (2008), 11-22.
Last retrieved from <http://www.iro.umontreal.ca/~blanchae/CATS2008/CATS2008.pdf> April 10, 2009.
- [2] R. Mizoguchi, E. Sunagawa, K. Kozaki, & Y. Kitamura, A model of roles within an ontology development tool: HOZO. *Journal of Applied Ontology* 2(2) (2007), 159-179.
- [3] N. Guarino, C. Masolo, A. Oltramari, L. Schneider, Sweetening Ontologies with DOLCE. In Proceedings of the 13th International Conference on Knowledge Engineering and Knowledge Management (EKAW02), *Lecture Notes in Computer Science* 473, (2002).
- [4] S. Alvino, P. Forcheri, M.G. Ierardi, & L. Sarti, A general and flexible model for the pedagogical description of learning objects, *Proceedings of WCC2008*, Milano, Italy, September (2008).
- [5] M. Grandbastien, F. Azouaou, C. Desmoulin, R. Faerber, D. Lecllet, & C. Quénu-Joiron, Sharing an ontology in education: lessons learnt from the OURAL project, *Proceedings of Seventh IEEE International Conference on Advanced Learning Technologies (ICALT 2007)*, Niigata, Japan, July (2007). Last retrieved from <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=04281129> on April 10, 2009.
- [6] D. Paneva, Ontology-based student modeling, *Proceedings of the Fourth CHIRON Open Workshop Ubiquitous Learning Challenges: Design, Experiments and Context Aware Ubiquitous Learning*, Turin, Italy, September (2006), 17-25.
- [7] L. Razmerita, A. Angehrn, T. Nabeth, On the role of Actor models and Actor modelling in Knowledge Management Systems, in *Proceedings of HCI International Conference*, Greece (2003).
- [8] D.A. Wiley, Connecting learning objects to instructional design theory: a definition, a metaphor, and a taxonomy. In: D.A. Wiley (ed.), *The Instructional Use of Learning Objects*, Association for Instructional Technology, (2000). Last retrieved from <http://reusability.org/read/chapters/wiley.doc> on April 10, 2009.
- [9] R. Andersson, A.K.Olsson, Fields of Education and Training, *EUROSTAT* (1999).
- [10] G. Paquette, An Ontology and a Software Framework for Competency Modeling and Management, *Educational Technology & Society* 10(3) (2007), 1-21.