Student modeling services for hybrid web applications

Iván Cruces[1], Mónica Trella[1], Ricardo Conejo[1], Jaime Gálvez[1]

[1] Dpto. Lenguajes y Ciencias de la Computación.
Universidad de Málaga 29071 - Málaga (Spain)
{icruces, trella, conejo, jgalvez}@lcc.uma.es

Abstract. This paper introduces a set of resources that provide web learning environments with student modeling services. SAMUEL is a user modeling server for registering, updating and maintaining student knowledge data from different sources that use their own ontologies. In order to make inferences about student knowledge, it becomes necessary to establish equivalences between concepts of different domains. For this purpose we have developed SONIA, a tool to edit and integrate models which allows semi-automatic ontology mapping. The set of resources is completed with INGRID, an independent generic open learner model for interacting with external student models.

Keywords: Web Intelligent Learning Environment, Student Modeling Server, Open Learner Model

1 Introduction

In the last decade, several adaptive and intelligent web-based learning environments (WILE) have been developed. Most of these systems are the result of research efforts focused on a particular pedagogical task, learning domain or teaching strategy, like for example ELM-ART [1] or AlgeBrain Tutor [2] which deal with LISP and Algebra domains respectively. Although less common, there are also educational web-based tools for generic domains such as SIETTE [3] or DCG [4]. Given the availability of all these tools, educators and course designers may be interested in integrating some of them in their own courses. Since these and other systems are high-quality software based on solid theoretical foundations, they may be of great value for the development of other educational systems. However, modularity is limited in most cases, which makes reusability almost impossible unless the system is used as a whole.

When thinking about developing tools that allow reusability of existing components, desirable features for such tools are: domain-independence (tools can be used for any subject domain); extensibility (in the sense that any component can be integrated), and component interoperability (components can communicate and interoperate despite differences of implementation language, execution environment

[1] This work has been partially supported by the ARTEMISA project number TIN 07-67515 of the MINISTERIO DE CIENCIA E INNOVACIÓN, Spain.

or model abstraction). Interoperability can be approached from two different angles: as a distributed software problem, and as a semantic conceptual issue.

Nowadays, web-based learning environments are evolving by adapting their architectures to new Web 2.0 technologies. The new approaches rely on the development of distributed architectures based on the integration and reuse of learning activities. Consequently WILEs become hybrid web applications (mashup).

Furthermore, the web teaching/learning model evolves towards an auto-learning scenario in which students complete their training using resources located all over the web. In any case, we deal with students who are using different systems. If these systems are intelligent (i. e. manage their own student model), each one of them stores part of the information about the student's knowledge, that is, the student model is distributed over the network. In this context, it would be useful that a system "asks someone for references" when it has to work with a new student, in order to provide adapted instruction.

This problem can be approached by using user modeling servers (UMS) [5;6;7]. According to Kobsa [8], *the purpose of user modelling servers is to separate user modelling functionality from user-adaptive application systems. They are not a part of an application system but rather independent from it*. In this way, an UMS is part of a distributed learning environment which provides teachers and users with educational services. $ADAPT^2$ [9] is an example of a framework for distributed education that integrates an UMS as part of its architecture. It stores students' activity and infers student's knowledge. When a student interacts with systems that have their own UMS, $ADAPT^2$ facilitates user models integration by means of Ontology Servers. It allows the knowledge of the same student to be modeled by multiple systems along different ontologies and stored on different ontology servers. However, once several adaptive systems decide to collaborate in sharing and exchanging student models, they have to select one specific ontology ;then the server used for user model exchange will be the one that hosts the selected ontology.

Our research group have been working during the last years in MEDEA [10], a distributed framework for the integration of web-based educational systems. MEDEA provides authors with the core components for an intelligent learning system: domain, student and instructional modules. MEDEA provides students with the necessary guidance during the learning process. It decides at each moment the most adequate task to be performed by the student. These tasks are performed by external learning resources (LR). Early version of MEDEA uses a central ontology and allows to manually map the domain model of LR into it. So, the user information received from LR could be translated into the concepts of MEDEA's ontology in order to update the user model. One of the main weaknesses of this version is that LRs cannot share data among them. They can only update the central user model but none other system can take advantage of it. Second, semantic integration of external adaptive systems depends on a manual mapping which is a costly task that needs the intervention of teachers and domain experts. In order to address these issues, MEDEA has evolved toward a more decentralized architecture. A set of resources that provides web learning environments with students modeling services has been developed and is presented in this paper.

SAMUEL (Spanish acronym for User Modeling Accumulative Server for E-Learning), is a User Model which allows storing student knowledge evidences obtained from different learning resources. SAMUEL is an independent component that can be requested by any web LR that needs to obtain data about a certain student.

SAMUEL stores heterogeneous information from different systems that deal with different domain ontologies. In order to perform evidence integration it is necessary that systems agree on the semantic of the domain terms, so they can exchange data for the equivalent domain concepts to update their own models.

Some attempts can be found in the literature for user model integration. In [11] authors propose a conversational model for reaching an agreement over not shared concepts. An ontology-based approach is used in [12] to identify similar concepts in the ontologies of related domains and align the domain models of two adaptive educational systems. In this paper we present an initial proposal based on ontology mapping techniques [13]. For this purpose SONIA (Spanish acronym for Intelligent Ontology Server for E-Learning) has been developed. This tool allows domain model edition and semi-automatic ontology mapping. SONIA stores the domain ontologies of external LR. When a student's mark in a domain concept is requested from Samuel, SONIA provides a list of equivalent domain terms. Then evidence integration heuristics are applied to all the concepts included in the list.

The set of student modeling services is completed with INGRID (Spanish acronym for Domain Independent Graphic Interface). This system allows students to consult (via Web) user models of resources with which they work and interact.

In the next section a description of different usage scenarios of MEDEA is discussed. In the rest of the paper, each tool is described in detail: SAMUEL (user modeling server) in section 3; SONIA (ontology mapping tool) in section 4, and INGRID (Open Learner Model) in section 5. Finally some conclusions of this work are presented.

2 MEDEA scenarios

In this section we describe new MEDEA architecture through different usage scenarios.

MEDEA provides all the components needed to create and execute an Intelligent Learning Environment: support for Domain Model definition, User Model (data storage and diagnosis processes), Instructional Planner and User Interface. All of them can be used together to act as a WILE while some of them can be used independently by any Educational Software.

The figure 1 shows the scenario 1, where all MEDEA components are used (MEDEA planner isn´t introduced in this paper). Teachers create the domain model and decide which learning resources are adequate to each concept (dotted line). Then a student interacts with the WILE through a web browser as described below.

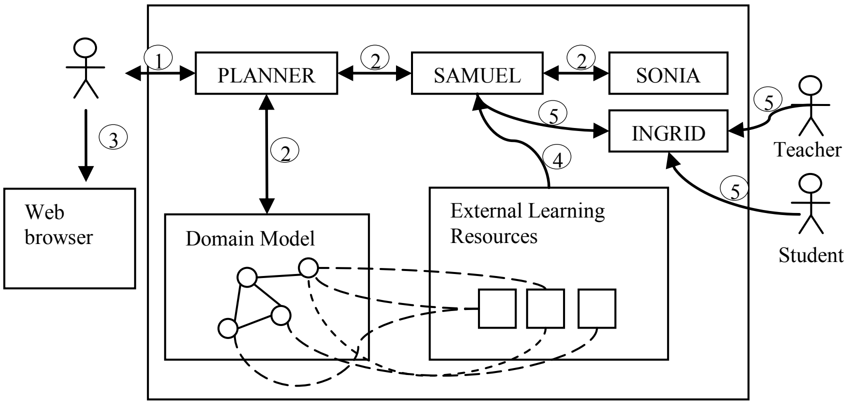

Fig. 1. MEDEA architecture - Scenario 1

Scenario 1

1. Student asks MEDEA planner for advice.
2. MEDEA planner consults domain model and student model (SAMUEL). SAMUEL uses SONIA to obtain all the data related to MEDEA domain model concepts. Each domain topic is related to one or more external LR that can be used to learn it. MEDEA planner suggests next topic to be learnt and some tasks (LR) to be done by the student..
3. Student works with a LR.
4. LR updates student model invoking SAMUEL web services.
5. Teachers and/or students can consult the student model using INGRID (Graphic Interface).

Besides being used together, each resource can constitute an independent component which may be used in combination with an LR. MEDEA has therefore evolved into a mashup. The figure 2 shows another possible scenario where some MEDEA components are used by external resources.

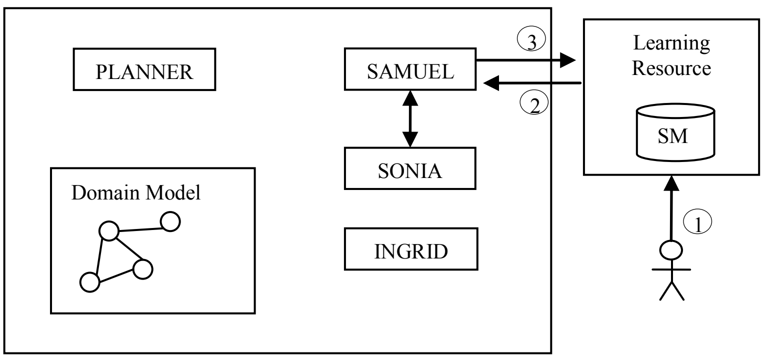

Fig. 2. MEDEA architecture – Scenario 2

Scenario 2

1. Students works independently with a LR .
2. The LR asks SAMUEL for information about this student's knowledge of a concrete topic. SAMUEL, using SONIA, compiles all the available data and sends it to the LR.
3. When the student finishes, the LR updates its own student model and can update SAMUEL too.

The figure 3 describes how INGRID can be used by a LR that has no way to show graphically its student data.

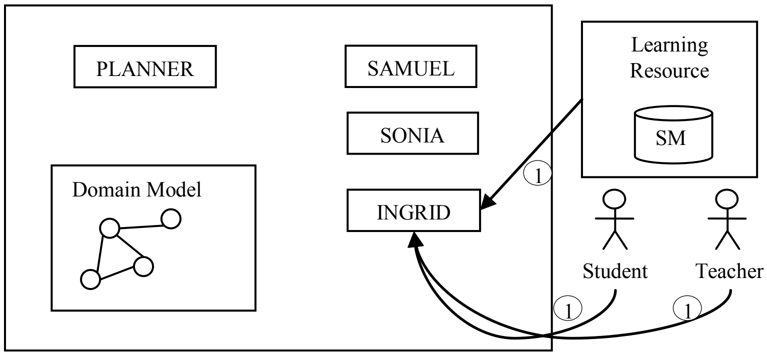

Fig. 3. MEDEA architecture - Scenario 3

Scenario 3

1. An external LR uses INGRID as a display for its model.

3 User model server for e-learning

SAMUEL is a user modeling server that allows storing of student knowledge evidences about different domain concepts obtained from different learning resources (Figure 4). Each LR can manage its own domain model, so each topic (domain concept) mark is stored independently together with the student ID, the domain and the source LR Therefore it is necessary, in order to make inferences about student knowledge, to establish equivalences between concepts of the same domain that belong to a different ontology. For this purpose we have developed an Intelligent Ontology Server for E-Learning (see next section).

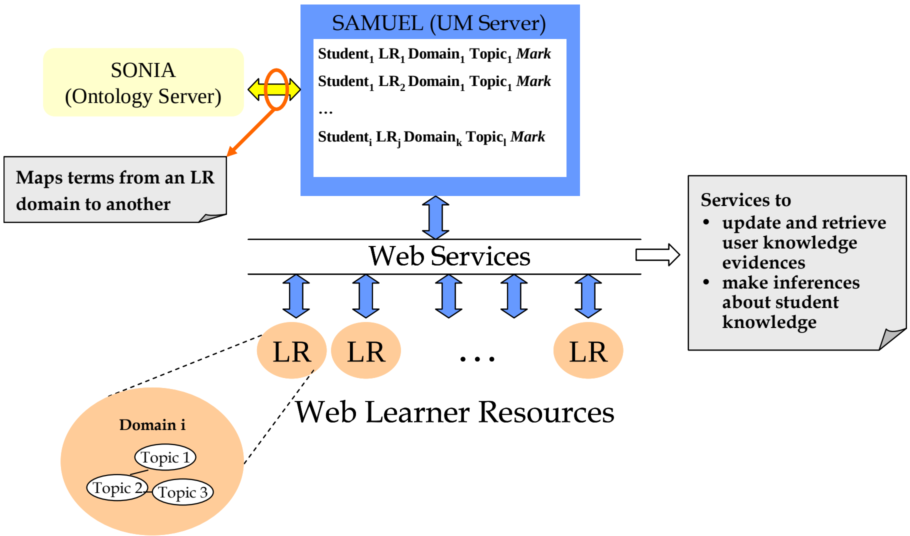

Fig. 4. SAMUEL architecture

For example, if we consider two WILEs which have a domain related to economics, we could map the terms as in Table 1.

Table 1. Economic domain ontology.

WILE 1
Topic 1: Economic concepts Topic 2: Kinds of markets
WILE 2
Topic 1: Introduction Topic 2: Financial system Topic 3: Markets …
Maps terms
(WILE 1) Topic 1: Economic concepts < - > (WILE 2) Topic 1: Introduction (WILE 1) Topic 2: Kinds of markets < - > (WILE 2) Topic 3: Markets …

If a student performs a task with WILE 1 about *Kinds of markets* topic and a task with WILE 2 about *Markets* topic, both, WILE 1 and WILE 2, will register the evidences in SAMUEL. So, if WILE 1 requests the mark of student in *Kinds of markets* topic, SAMUEL will infer it with *Kinds of markets* evidences and *Markets* evidences, through the mapping terms.

At present SAMUEL offers a set of web services implemented with JAX-WS API, which allows web learning resources to register evidences and to obtain data about other user activities. Besides a concept mark, each record contains information about the learning activity that provides the information. The services have three main parameters: the set of evidences used to estimate the student's knowledge, the evidence sources and the method used to obtain the estimation.

Table 2. Parameter values of SAMUEL services.

Nº evidences	Source	Estimation
Last evidence	Specific source	Average mark
Last n evidences	A set of sources	Weighted mark
Evidences in a date range		

According to the values showed in Table 2, SAMUEL can be asked to return the *last n* evidences concerning a concept or those collected in a period of time. In both cases, clients can retrieve average and weighted marks for any concept. These data can be referred to one or more learning resources. The heuristics used to calculate the weighted marks are represented by equation **1**.

$$mark(c) = \frac{\sum_{i=1}^{n} w_i m_i}{\sum_{i=1}^{n} w_i} \quad \textbf{(1)}$$

In *Evidences in a date range* case the *n* value in equation **1** represents the total number of days in the date range and the m_i value is the average mark of day i. In *Last n evidences* case the *n* value represents the n parameter and the m_i value is the mark of the *i-th* last evidence registered. The w_i values are the mark weights. One example could be $w_i=1/i$, where more weight is given to the most recent evidences since they represent more accurately the student's current knowledge.

4 SONIA

As it has already been discussed, in order to make inferences about student knowledge using data collected from different learning resources, it becomes necessary to establish equivalences between concepts of different domains. We have approached this problem with the development of a web ontology server called SONIA. It is an AJAX application that allows editing concept semantic networks (first level ontologies) and establishing relations among them semi-automatically (Figure 6).

Stemming algorithms (for English and Spanish languages) are applied to ontology terms and then a set of *string similarity metrics* (Hamming distance, N-gram, Levenshtein distance and Maximum common substring) are used to calculate the probability of two terms from different ontologies referring the same concept. In addition to these metrics, a synonyms dictionary has been added.

The Figure 5 shows an example where terms of two ontologies are compared. Each arrow's color represents a different metric. The number above the arrow indicates the similarity degree between the terms. These correspondences can be updated (inserted and/or deleted) manually.

These mappings are stored and, when SAMUEL is requested information about a concept (Figure 4), it searches for all the terms related in order to collect data which refer to the same concept and have been provided by different WILEs.

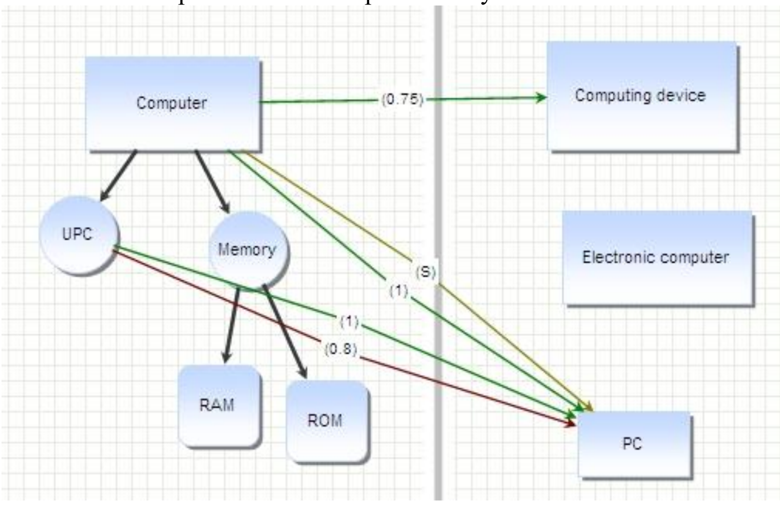

Fig. 5. Ontology mapping in SONIA.

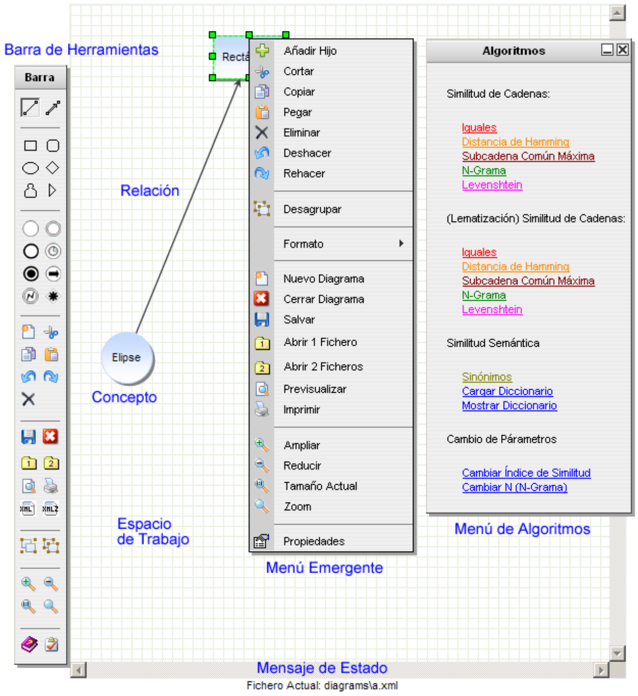

Fig. 6. SONIA interface.

5 Domain independent graphic interface for student models

Open learner models (OLM) are accessible and open models that extend the traditional WILE models to turn them into a visible and interactive part of the system [14]. OLM allows a student to inspect his model and interact (edit or negotiate) with it. This kind of system stimulates the student's analytical thinking and helps him/her to plan and monitor his/her learning [15]. OLMs are not just valuable for students but also for teachers. In fact, a graphical representation of the model can help teachers to carry out a course formative evaluation and to determine the students' learning problems. Moreover, the model accuracy can be improved if the system allows students and teachers to collaborate in the modeling process. Broadly speaking, the student model has evolved from being a knowledge source for learning resources (closed system) to become an important learning resource for the student (open systems).

There are two types of OLM: those integrated in a tutor system and the independent ones (IOLM), whose purpose is to help students to identify and to solve learning problems by themselves, without tutor system help, that is, to encourage metacognitive skills.

Still in the field of resource integration, we have developed INGRID [16], a web IOLM which allows students to consult the user models of the different learning resources with which they work and interact. INGRID provides two views of student score, both based on the relationship topic / sub-topic from the domain concepts. The first is a hierarchical structure (Figure 7) representing the issues by a graph, and the second is a table of topics and marks. The hierarchical view of the graph represents the nodes with a color code indicating the student's level of knowledge in this concept according to a particular source (SIETTE system, for example). The table of topics shows bars that represent the marks on a scale from one to ten. Moreover, each topic can have several actions associated to it to edit the user model (e.g. SIETTE tests).

The strengths of the system are that it is generic and it can represent data from any WILE. For this purpose, INGRID has a JSP which receives as input (xml format) a list of concepts and the marks obtained by the students, as well as the semantic network of the domain model (concepts and relationships) and is capable of representing it. So far it has been successfully tested with SIETTE and user model server SAMUEL. Both tools use it as a plug-in to graph the data from their students.

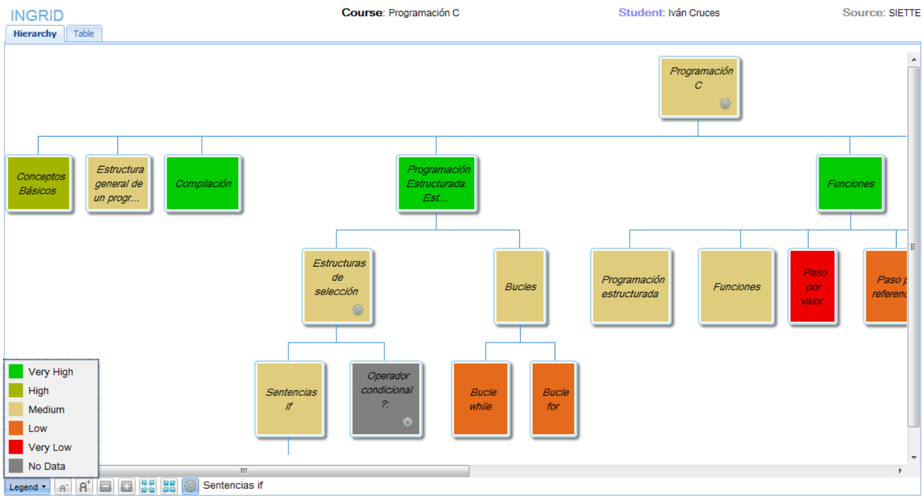

Fig. 7. Hierarchical view in INGRID.

6 Conclusions

The general objective of our research work is the development of different domain independent and interoperable components that can be used in the construction of a web based intelligent learning management systems.

The user interaction with different learning resources can provide valuable information that helps other systems to provide the student with a more accurate instruction. In this sense we work in the integration of heterogeneous sources of user information. In order to achieve this, our group is working on the development of independent and generic tools that provide student modeling services. A functional approach has been presented in this paper. It includes a set of tools that can be used in any learning environment to share user data.

In this first stage of the work, all the tools needed to provide user modeling services have been implemented and all of them are working successfully in a real environment. As part of a formative evaluation process, these tools have been integrated with the test system SIETTE [17]. At present, we are gathering user data from SIETTE, and we are planning to use new learning resources as data sources in the near future. The results obtained from this evaluation have opened up possibilities for future research.

First, we are aware of the importance of the semantic issue in WILE interoperability. It is a bottleneck in intelligent learning resources integration. A first approach to this problem is SONIA. So far, we have applied ontology mapping techniques based on lexical components. The next step is to use techniques based on ontology structure as graph matching and studying relations semantic. In the next stage of this work we plan to conduct research into the application of machine learning or statistical techniques for ontology mapping.

Furthermore, we have implemented a set of services to accumulate user evidences. We plan to add new services which apply formal diagnosis methods to make inferences about student.

Finally, we think that teachers and students are an important source of evidences, therefore we will not only allow INGRID to consult the server data but also to update them taking into account teachers and students' contributions.

References

1. Weber, G. and Brusilovsky, P.: ELM-ART: An adaptive versatile system for Web-based instruction. Special Issue on Adaptive and Intelligent Web-based Educational Systems. International Journal of Artificial Intelligence in Education, vol. 12, pp. 351--384 (2001)
2. Alpert, S. R., Singley, M. K., and Fairweather, P. G.: Deploying Intelligent Tutors on the Web: An Architecture and a Example. International Journal of Artificial Intelligence in Education , vol. 10, pp. 183--197 (1999)
3. Conejo, R., Guzmán, E., Millán, E., Pérez-de-la-Cruz, J. L., and Trella, M.: SIETTE: A Web-Based Tool for Adaptive Testing. International Journal of Artificial Intelligence in Education, vol. 14, pp. 29--62 (2004)
4. Vassileva, J. and Deters, R.: Dynamic courseware generation on the WWW. British Journal of Educational Technology, vol. 29, pp. 5--14 (1998)
5. Kay, J., Kummerfeld, B., and Lauder, P.: Personis: A server for user modeling. Proceedings of Second International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems (AH'2002), 201--212. Málaga, Spain (2002)
6. Brusilovsky, P., Sosnovsky, S., and Shcherbinina, O.: User Modeling in a Distributed E-Learning Architecture. In: L. Ardissono, P. Brna and A. Mitrovic (eds.) Proceedings of 10th International User Modeling Conference (Edinburgh, UK, July 24-29, 2005). Lecture Notes in Artificial Intelligence, vol. 3538, pp. 387--391. Springer-Verlag, Heidelberg (2005)
7. Kobsa, A., & Fink, J. (2006). An LDAP-based User Modeling Server and its Evaluation. User Modeling and User-Adapted Interaction 16(2), 129-169.

8. Kobsa, A.: Generic User Modeling Systems, in The Adaptive Web. LNCS, vol. 4321, pp. 136--154. Springer-Verlag, Heidelberg (2007).
9. Brusilovsky, P., Sosnovsky, S., and Yudelson, M. (2005) Ontology-based framework for user model interoperability in distributed learning environments. In: G. Richards (ed.) Proceedings of World Conference on E-Learning, E-Learn 2005, Vancouver, Canada, October 24-28, 2005, AACE, pp. 2851-2855.
10. Trella, M., Conejo, R., Guzmán, E. y Bueno, D. (2003): An Educational Component Based Framework for Web ITS development. Web Engineering. International Conference, ICWE 2003. LNCS, vol. 2722, pp. 134--143. Springer-Verlag, Heidelberg
11. Cena, F., Furnari, R. (2008): A soa-based framework to support user model interoperability. In: Nejdl, W., Kay, J., Pu, P., Herder, E. (eds.) AH 2008. LNCS, vol. 5149, pp. 284–287. Springer, Heidelberg
12. Sosnovsky, S., Dolog, P., Henze, N., Brusilovsky, P., & Nejdl, W. (2007). Translation of overlay models of student knowledge for relative domains based on domain ontology mapping. In: R. Luckin, K. R. Koedinger, & J. Greer (eds.) *Proceedings of the 13th International Conference on Artificial Intelligence in Education (AIED'2007), Marina Del Ray, CA, USA, July 9-13, 2007*, 289-296.
13. Y. Kalfoglou, M. Schorlemmer (2003). Ontology Mapping: The State of the Art.
14. The Knowledge Engineering Review 18 (1):1-31, 2003.Bull, S., Kay, J. (2007): Student Models that Invite the Learner In: The SMILI:() Open Learner Modelling Framework. International Journal of Artificial Intelligence in Education, vol. 17, pp. 89--120
15. Mitrovic, A., Martin, B. (2002): Evaluating the Effects of Open Student Models on Learning. LNCS, vol. 2347, pp. 296--305. Springer-Verlag, Heidelberg
16. INGRID, http://marte.lcc.uma.es/ingrid/index.jsp&demo=1
17. SIETTE, http://www.siette.org
18. MEDEA, http://marte.lcc.uma.es/medea