Enrichment of Learner Profile with Ubiquitous User Model Interoperability

María de Lourdes Martínez-Villaseñor[1], Miguel González-Mendoza[2], and Ignacio Danvila Del Valle[3]

[1] Universidad Panamericana, Campus México, D.F.,
Mexico

[2] Tecnológico de Monterrey, Campus Estado de México, Edo. de México,
Mexico

[3] Universidad Complutense de Madrid, Madrid,
Spain

lmartine@up.edu.mx, mgonza@itesm.mx, idanvila@ccee.ucm.es

Abstract. Nowadays, there is a constant need of acquiring new knowledge and skills to keep up with the demands of changing environment. The design and development of training and educational systems that enable effective personalized learning help obtaining changing skills and fill competence gaps. The computational effort to create a user model that represents user's knowledge, characteristics, interests, goals, background and preferences is repeatedly done by many systems and applications in several domains. Each system ends up with a partial view of the user. Researchers in user modeling foresee the need of sharing and reusing user model information in order to obtain a better understanding of the user and be able to provide personalized and proactive services. In this paper we present an application scenario of sharing and reusing information scattered in most commonly used applications to enhance learner profiles.

Keywords. User modeling interoperability, learner profile enrichment.

Enriquecimiento del perfil del estudiante mediante la interoperabilidad de modelos de usuario ubicuos

Resumen. En la actualidad hay una necesidad constante de adquirir nuevo conocimiento y habilidades para cubrir las demandas de un ambiente cambiante. El diseño y desarrollo de sistemas educacionales y de entrenamiento que permitan un aprendizaje efectivo y personalizado, ayuda a obtener habilidades cambiantes y llenar las brechas de competencia. El esfuerzo computacional para crear un modelo de usuario que represente el conocimiento, características, intereses, metas, antecedentes y preferencias del usuario es realizado repetidamente por varios sistemas y aplicaciones de diferentes dominios. Cada sistema termina con un conocimiento parcial del usuario. Investigadores del área de modelado de usuario visualizan la necesidad de compartir y reusar la información de los modelos de usuario para obtener un mejor entendimiento del usuario y proveer servicios de manera personalizada y proactiva. En este artículo presentamos un escenario de aplicación para compartir y reusar información esparcida en las aplicaciones más comúnmente usadas con el fin de enriquecer perfiles del estudiante.

Palabras clave. Interoperabilidad de modelos de usuario, enriquecimiento del perfil de estudiante.

1 Introduction

In different domains, user-adaptive systems adjust their contents, structures and interfaces to certain features gathered in user models [1]. A user model is essential to enable a system to behave differently for distinct users [2]. So the quality of the services delivered depends on the accuracy and coverage of the user model. A better knowledge of the user allows the adaptive systems to provide more personalized and proactive web applications and services delivered with ubiquitous access.

Depending of the purpose and goals of a system, different information about the user is needed. In the educational systems domain, computer systems have gathered learner information about employees, pupils and students since the seventies [3]. The user model is known as a *student model* for Intelligent Tutorial Systems or *learner model* in other educational systems. Often, a learner model is a partial view of the user, an overlay of the domain model that tries to represent his/her current state with respect to the domain concerned [4].

Many learning platforms do not take into account the different needs of the learner and provide the same information retrieval and learning content to every user. The use of learner model enables effective personalization of learning environments. Adaptive and personalized learning is particularly vital nowadays, when skills become outdated rapidly demanding constant training for higher skilled jobs.

In any adaptive educational system, information of the user (learner) is gathered from explicit information provided by the user, and/or implicit information obtained from the observation of user's interaction and behavior. The computational effort of building a model is often repeated across different platforms, applications and services in various domains. Likewise, the user has to invest time in setting parameters of devices and services repeatedly which leads to overabundance of information, inconsistencies, redundancies and repeated configurations. Researches in many domains [5, 6] as well as learning environments [7] envision the need to share and reuse information of user models and learning information in order to ease the pain of building a user model from scratch and deal with the "cold start" problem.

With a better understanding of the user, adaptive systems can provide a better service, content and interface personalization and adaptation. Sharing and reusing information between models can bring advantages for profile providers and profile consumers; it helps dealing with the above mentioned problems, provides enrichment for the user models and prevents the user from repeating the process of configuration.

Gathering distributed user information from heterogeneous sources implies interoperability problems [6]. "In modern information environment there is a constant increase of amount and complexity of data being stored and used" [8]. The integration of distributed user information from heterogeneous sources and making sense of it to enable user modeling interoperability, entails handling the syntactic and semantic heterogeneity of the user models [9]. Regarding learning information interoperability Paulsson [3] finds that policies, proprietary formats and lack of exchange mechanisms hinder learner information interoperability. In order to enable interoperability and scalability of learner profiles, an overwhelming number of standards and specifications have been proposed. Some important examples are IMS Learner Information Package (LIP) [10] and IEEE-LTSC formerly called Personal and Private Information (PAPI). Although these standards can simplify the exchange of information across distributed computing environments, they are not commonly adopted. These standards hinder the interoperability process between learner profiles because "learning standards are not harmonized to work with each other" [11]. The lack of sharing mechanisms and low capability of extensions hinders sharing and reusing learner profiles [11].

We discuss how to take advantage of profile information from distributed heterogeneous sources to enrich learner information of adaptive educational systems. People interact nowadays with systems and applications through several devices. Each user end up with valuable information about his/her preferences, interests, goals, knowledge and characteristics scattered in distributed heterogeneous user models. Each application has just a narrow understanding of the user, given by some features of the user's profile. In previous works [12, 13] we presented a framework for ubiquitous user model interoperability that enables sharing and reusing user's profile information from commonly used sources like social network applications, FOAF, personal health records and devices.

In this paper, we present an application scenario of reusing and sharing user profile information from distributed heterogeneous sources to enrich learner profile through the mediation of our ubiquitous user model. Our framework helps reuse valuable information

scattered in distributed heterogeneous sources of diverse domains and enhance the learner profile of adaptive educational systems with the least intervention of the profile providers.

The rest of the paper is organized as follows. In Section 2, we discuss the related work in ubiquitous user model interoperability and briefly describe our on-going work regarding this problem. In Section 3, we describe the application scenario of enrichment of learner profile information using our framework to enable ubiquitous user model interoperability. In Section 4, we present a proof-of-concept demonstration and first results of learner profile enrichment using Moodle basic profile as an example. Our conclusions and future work are described in Section 5.

2 Ubiquitous User Model Interoperability

Current research in ubiquitous user modeling has two major approaches: (i) standardization based user modeling based in semantic standardization of user model defining some common ontology and language; (ii) mediation-based user modeling using mediation techniques to build semantic bridges between representations. Recent surveys about user modeling interoperability [5, 6] indicate that adhering to a standard ontology and representation is not feasible in a distributed open environment, and mediation approach means dealing with syntactic and semantic heterogeneity and many transfer mechanisms. Berkovsky *et al.* [5] suggest bridging the gap between these two approaches.

In this section, we review representative research literature of major user model interoperability approaches in learning environment taking into account the representation of the exchanged data. We also present a summary of our framework for ubiquitous user model interoperability.

2.1 Standardization-based User Modeling

Standardization-based user modeling approaches try to deal with syntactic and semantic heterogeneity of different learner profile providers

and consumers using one or more interchange standards. Standards are basic common means of understanding to enable interoperability, but educational and training systems rarely adhere completely to a standard. Systems usually need some particular information that is not contemplated in any standard. Therefore, researchers aim to leverage standards to gain more flexibility for their learner representations.

Dolog and Schafer [14] described a framework based on IMS LIP and IEEE PAPI to enable exchange learner profiles between e-Learning and learner assessment systems. User modeling systems interested in learner profile exchange have to provide mappings between their internal data models and the author´s standard based descriptions. Musa and De Oliveira [15] proposed a Web Service based architecture to share learner´s information between e-learning systems. The authors created a learner model joining categories of two well-known standards, IMS LIP and IEEE PAPI, in order to enrich the model and enable learner information interoperability. They provide additional definitions of learning styles and cognitive styles. These solutions allow learner information exchange and improve the aforementioned standards, but they still lack necessary flexibility for dynamic environments.

Brusilovsky *et al.* [16] presented KnowledgeTree, an architecture that enables reusing learning activities in E-learning environments. In this paper, the authors described CUMULATE, a student model server that collects evidence of student activity from several servers. This architecture has problems to integrate components that have their own user models and user functionality. In [17], the authors presented a general ontology-based general framework for distributed user model interoperability of adapted Web-based systems (ADAPT). The extended ADAPT architecture includes an ontology server containing centralized metadata acting as a common storage and representation of student's knowledge. The user model exchange is possible if there is an agreement to use a specific ontology.

Another example of common user model approach is Personis AD [18]. Personis [19–22] includes variations of an ontology-based reusable and understandable student model. The structure

of the user models was based on a set of predefined and agreed upon ontologies enabling the accretion of user models and focus in the scrutability and privacy issues. Personis AD [18] is an evolution of Personis for context-aware ubiquitous applications which supports distributed models and associated resource discovery.

The authors of [12] presented an open and generic learner based on two standards: IMS LIP and CIM User Model [23]. This open learner profile facilitates sharing and reusing learner information of hosted applications implemented in a WBEM (Web-Based Enterprise Management) architecture. They provide two management services to modify and extend the core learner schema and manage the learner profile repository. This work is a representative example of learner information interoperability done with a standardized-based approach which implies adhering to certain standards in order to be able to participate in the interoperability process being managed by a centralized server.

2.2 Mediation-based User Modeling

In the mediation or translation user modeling approach, applications have their own representation of the user model and mapping has to be provided in order to align the different models. Multiple syntactic and semantic conflicts must be resolved. Providing mappings between user or learner models is not an easy task and it is frequently done manually, using generic schema, ontology matching tools.

[23-27] suggested a distributed multi-agent approach to cope with fragmented and inconsistent user models in ubiquitous computing and achieve the integration of user models to deliver personalized services. In the previously mentioned works, the authors presented i-Help in which broker-agents keep track of user models and are able to map help requests to possible service providers based on domain taxonomy.

One example of interoperability between Adaptive Educational Hypermedia is presented in [28], specifically between user models of MOT and WHURLE. Although this work extracted commonalities between AEH systems, the conversion analysis was made by human experts. MEDEA [29] is an open learning platform for the

development of intelligent web-based educational systems. MEDEA is considered as a centralized approach [6], because it provides a unified student model that summarizes the data of different types of instructional resources. The relationships between concepts are established by a human expert, so it is considered also as a translation approach. One representative example of mediation using semantic mappings is the Generic User Modeling Component (GUC) presented by Van der Sluijs and Houben [30] introducing Web 2.0 technology into user modeling servers. They used OWL for user model representation and schema matching techniques to deliver the appropriate user or learner model from the repository when a service request arrived. A human designer constructs a rule-based mapping schema. Mappings are not defined automatically.

In the domain of recommender systems, [31] aims to integrate multiple agent-based services in a unique representation Smart User Model (SUM) in distributed and heterogeneous environments using multi-agent technology. SUM is defined as a collection of attribute-value pairs representing objective, subjective and emotional features. The authors presume that the user features can be learned from user information extracted from distributed user models and deliver them to other recommender systems. Two examples of reusing profile information from different recommender systems in the prediction generation process are [32, 33]. Berkovsky *et al.* [34] proposed a general framework and methodologies for enhancing the accuracy of user modeling in recommender systems importing and integrating data collected from different recommender systems. In this work, the authors addressed the challenges of sharing user modeling information in heterogeneous sources with the defined process of *user models mediation*. Models are stored in a decentralized way in the service providers' side, but mediation is done through a centralized point of access. In several papers [35-37], Berkovsky shows examples of the four different kinds of mediation described in [34]: cross-representation, cross-user, cross-item and cross-context mediation.

More recently, Carmagnola *et al.* [9] presented a solution with high flexibility representing user models and providing semantic mapping of user

data from heterogeneous sources. However, to take part in the interoperability process, every provider must comply with a standard format for exchange and maintain a sharable user model which includes fragments of it as RDF statements. Sosnovsky *et al.* [38] presented an infrastructure to collect and integrate evidence about student knowledge of different distributed student models. Their approach is based on manual mappings of different domain student models which allow a student to use two adaptive educational systems in parallel.

2.3 Bridging the Gap between Mediation and Standardized Semantic Integration

Nowadays, in ubiquitous environment, everything changes:

–   Profile providers and consumers constantly appear and disappear in the user modeling interoperability scenario.
–   User models are not static, they constantly evolve. The user make explicit changes in profile information; from user's usage and behavior, new knowledge can be implicitly acquired; user models structure itself can change over time.
–   Context of the user always changes regarding time, space, devices used, preferences, among other factors that change the user's situation and complicates the tasks of adaptation and personalization.
–   User models are being built in many domains.
–   The users of applications are not static; they come and go using different services.

The above mentioned facts imply the following problems when trying to share and reuse profile information:

–   Great syntactic and semantic heterogeneity in the profile suppliers and consumers which hampers communication between adaptive systems.
–   It is difficult to define a standard, because all static attempts cannot cope with a dynamic ubiquitous environment. Nevertheless, some semantic web standards have proven to be very useful as a start point of understanding between heterogeneous sources. In

summary, a ubiquitous user model must be able to evolve over time, but some basic or standard profile and format help to ease the communication.

–   If mediation is done without a central user model, the addition of a new stakeholder in the process of user modeling interoperability entails creating two transfer mechanisms, two semantic mappings and probably two translation functions for each existing profile supplier or consumer. The complexity of this task is from order of N^2 between every two applications.
–   When dealing with multiple user and domain models possible conflicts may occur [6, 38]: naming conflicts, different graph structures, different scopes, different granularity, and different focus. This conflicts need to be resolved when the task of model alignment is done.
–   Contextual and privacy issues complicate the tasks of interpretation and use of user modeling data.

A possible solution to overcome limitations of standardization and mediation approaches and leverage their advantages is to integrate elements of both approaches. Current state of the art solutions for ubiquitous user model interoperability must consider the high degree of dynamism of ubiquitous environments.

Some of the recent examples are trying to integrate information from different distributed sources in a central repository providing mediation between profile information providers and consumers [30, 34]. The authors of [39] proposed a learner profile model combining IMS LIP and FOAF (two well-known standards) adding context and agenda information, and multi-agent based mediation to develop pervasive learning environments. Their prototype collects information from autonomous social networking providers taking advantage of the valuable information provided in these frequently used applications. The authors of [40] defined a categorization of mapping types that highlights the difficulty of creating matches between elements of learner models. They proposed a framework that includes a learner mapping web application for the manual creation of mappings between educational web

systems. Manual mapping demands a lot of time and effort from experts, and falls short of what is needed to cope with highly changing environments.

Efforts have been made to deal with the high degree of dynamism and high levels of data heterogeneity, but we are not aware of fully implemented automatic semantic mapping between heterogeneous sources in the learning or any other environment. This would facilitate interoperability with any source at any time.

2.4 A Framework for Ubiquitous User Model Interoperability

As explained in the previous section, in a multi-application, multi-device ubiquitous environment, user profile information is scattered in distributed user models. When trying to integrate all this valuable user information, it is important to take into account that highly autonomous profile suppliers and consumers are participating in the interoperability process. This means, first of all, that providers are free to decide what data to store, how to describe the data, the set of constrains on the data, and associate an interpretation [41]. Providers also decide what data to share, policies and means of how to share it. Consumers of profile information want to decide when join and leave the system as well. Consumers have also their ways to describe and interpret data. Therefore mechanisms of interoperability must be provided that require the least intervention and effort of the ubiquitous user modeling stakeholders in order to enable interoperability respecting the providers' and consumers' autonomy. These conditions stand in learning environments. The user or learner information representation must be machine-readable and flexible to allow the integration of information of new providers.

In previous works [12, 13] we presented a framework for ubiquitous user interoperability. The proposed framework enables the interoperability between profile suppliers and consumers with a mixed approach that consists in central ubiquitous user model ontology to provide formal representation of the user profile and a process of concept alignment to automatically

discover the semantic mappings between the user models.

The central ubiquitous user model interoperability ontology (U2MIO) is a flexible representation of a ubiquitous user model to cope with the dynamicity of a distributed multi-application environment that provides mediation between profile suppliers and consumers. U2MIO can evolve over time to adapt the representation to the changing multi-application environment. The dynamic user profile structure ontology is based in Simple Knowledge Organization for the Web (SKOS) [42]

The process of concept alignment is briefly described below. This process automatically discovers the semantic mapping between the concepts of profile suppliers and consumers and the U2MIO ontology in order to interpret the information from heterogeneous sources, and integrate them into a ubiquitous user model. This process is crucial for the construction and maintenance of the ubiquitous user model; it enables interoperability and allows the evolution in time of the U2MIO ontology. We proposed a two-tier matching strategy for the process of concept alignment in a hybrid integration system to provide mediation between heterogeneous sources. This architecture and the process of concept alignment facilitate the participation of new stakeholders in the interoperability process.

Ubiquitous User Model Interoperability Ontology

The Ubiquitous User Modeling Interoperability Ontology (U2MIO) represents a flexible user model profile that evolves during time according to the recommendations of the concept alignment. The ontology reuses SKOS ontology designing a central concept scheme for the ubiquitous user model and one concept scheme for each profile supplier or consumer. Semantic mapping relations between each of the stakeholder's concept scheme concepts and the central user model concept scheme concepts are determined by the process of concept alignment. Semantic relations are set with SKOS properties. This representation supports interoperability overcoming semantic differences and enables the participation of new stakeholders in the interoperability process without effort of the profile
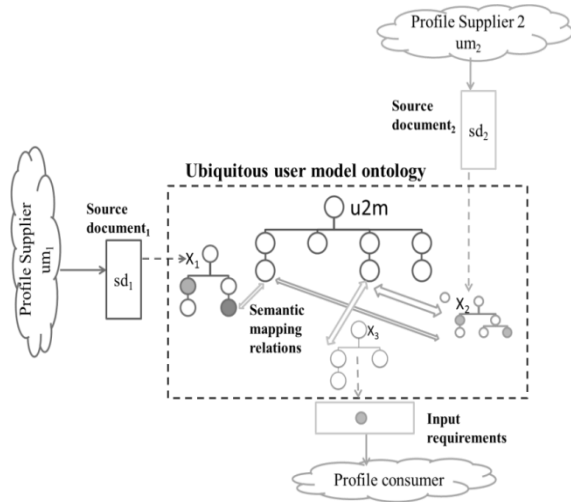
Fig. 1. Interrelations between profile stakeholders and ubiquitous user model
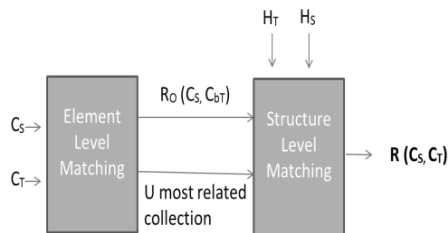

Fig. 2. Two-tier matching stategy of the process of concept alignment

information provider or consumer. Fig. 1 shows the interrelations between profile stakeholders and the ubiquitous user model concept.

Process of Concept Alignment

Our ubiquitous user modeling framework deals with the profile suppliers' transfer mechanisms and recollects source documents *(sd)* in XML, JSON or RDF. If the source is new to the system, a corresponding skos:ConceptScheme *(X)* is designed and added to U2MIO. The process of concept alignment is based in a two-tier matching strategy (Fig. 2). First, an element level matching step finds a set of concept candidates for alignment for each concept in the source concept scheme. This task is performed combining three types of similarity measures: a) string similarity based in Dice [43], b) a simple distance of longest

substring, c) semantic similarity based on WordNet [44]. From this step in which we analyze the word similarity between each concept in the source with all concepts in ubiquitous user model concept schema (*u2m)*, we find a set of best suited concepts for alignment (or one best suited concept) in the target (*u2m*). Next, the method looks for structure similarity. The goal in the structure level matching step is to disambiguate the meaning of the word analyzing its context, this means analyzing the structure and meaning of the neighbor concepts in the same source document. In this step, the similarity between the neighbor concepts in the source and the neighbors of the best suited concept(s) in the target are calculated. After this step, a set of IF THEN rules are applied to determine one-to-one semantic mappings and recommendations of concept and collection additions. The process of concept alignment shown in Fig. 2 roughly describes the inputs and outputs of each phase. A concept scheme is defined as *(C, H_C, V_C)* where *C* is a set of concepts arranged in a subsumption hierarchy H_C. V_C is the set of corresponding concept values. C_s is the set of concept labels extracted from the source document. C_T is the set of concept labels extracted from the target (ubiquitous user model scheme), and C_{bT} is the set of concepts that are best suited for alignment. $R_0 (C_s, C_{Tb})$ are the highest relations obtained from the element level matching phase and $R(C_s, C_T)$ are the final semantic mapping relations found between the concepts of the source document and the ubiquitous user model (target).

3 Application Scenario of Enrichment of Learner Information with Ubiquitous User Model Interoperability

Adaptive educational systems have to investigate different types of knowledge relevant to provide effective personalization of learning environments. The authors of [45] specify the following basic aspects for adaptation:

– The domain model that represents what is to be delivered and adapted.

– The user model and context model define parameters to select and adapt the domain content or presentation.

– The activity and adaptation model that represent how the delivery and adaptation should be performed.

Brusilovsky *et al.* [2] state that user model for adaptive hypermedia and adaptive educational systems frequently represents the user's knowledge, interest, goals background and individual treats. Information required depends on the class of adaptive system and adapting needs. Web-based adaptive educational systems (AES) focus in modeling knowledge and learning goals as Intelligent Tutoring Systems (ITS) do. Adaptive information systems and Web recommenders especially need to model interests, as in adaptive information retrieval systems. Adaptive hypermedia systems (AHS) require more information so they model knowledge, interests, goals, context of the work and individual traits [2]. Ubiquitous and mobile systems focus on context of the work and user's tasks and goals. Nowadays, learning systems collect information about the learner explicitly or implicitly. Systems and applications ask the users to insert information about their background, preferences, interests, and demographic features, or create a model based on learner's usage and behavior. Predefined learner stereotypes are frequently used to provide adaptation and personalization.

Although each of these adaptive educational systems focus on different aspects of the users, a better understanding of the learner enables to deliver tailored resources and services. "The more a learning system knows about the learner, the greater is the chance to deliver learning content that best suits to learner's needs" [46]. In order to achieve personalized adaptive learning, an adaptive system has to determine how to represent the user model, establish ways to communicate with the user model and gather user information. Each system and application repeats the cost and effort of building a user model, and the learner has to repeatedly give the same information in different systems. Sharing and reusing information of user models is the answer to deal with these problems and avoid having isolated sparse user information. Authors in [7]

discuss the state of the art and main challenges of interoperability in personalized adaptive learning. These challenges are very similar to user model interoperability in other domains presented in [6]. The interoperability process between distributed profile information providers and consumers implies dealing with syntactic and sematic heterogeneity.

For the purpose of sharing and exchanging user profile information, profile providers and consumers have to share a common representation and means of exchange or having a process of conversion between representations, as we reviewed in Section 2. In the adaptive educational systems scenario, learner profile standards, IEEE Personal and Private Information (PAPI) and the IMS Learner Information Package (LIP) try to establish a common representation and exchange format for diverse adaptive educational systems. Nevertheless, as standards are not universally adopted, conversion is still needed for some profile providers and consumers. The process of mapping one learner profile to another implies a lot of manual labor or in the best scenario, a complex semiautomatic process.

3.1 Contribution of Ubiquitous User Model Interoperability to Adaptive Educational Systems

As we explained above, there are many adaptive educational systems that model the user focusing in different aspects and features of the learner. Each system and application requires a lot of effort from the user to explicitly capture basic profile information, preferences and interest. They frequently need a lot of interaction from the user to implicitly observe her/his behavior and build a useful learner profile.

Some of these adaptive educational systems and standards have their own APIs and web services that enable populating profile information from other systems, but as they are autonomous regarding their learner model representation, means of communication and semantic meaning of the data, human interventions for mapping is commonly needed. Prior agreement and/or mapping demand a lot of human effort. Solutions that integrate and combine profile information still

demand human intervention to enable interoperability. Even using standards, agreement and/or manual or semiautomatic semantic mapping is needed.

On the other hand, there has been a great evolution of user modeling and adaptation from desktop paradigm to ubiquitous computing. People interact daily with systems and applications unrelated to educational domain through different personal devices explicitly sharing information about them or implicitly leaving trace of their behavior, preferences, and interests among other personal features. Valuable profile information of the user is gathered by systems and applications from diverse domains. Personal devices with attached sensors collect information that provides context of use, useful for personalization too. Social network applications, for instance, have become very successful and can be considered as valuable sources of personal information in order to build a user

model. According to Nielsen, in a study published in August 2012 [47], Facebook is the second more visited Web site by people in the U.S. (after Google). Semantic Web technologies such as FOAF and microformats represent an important step forward in terms of the Web's evolution. They provide a way to embed machine readable structured data into web pages for expressing user personal information [48]. FOAF can be an important profile supplier as it is currently considered as one of the best populated ontologies and is extensively adopted to describe users and their relations [49]. Therefore, RDF documents published using FOAF vocabulary are valuable sources for the user model.

In summary, valuable profile information is scattered in distributed heterogeneous sources of diverse domains that once gathered, crumbled at concept granularity, and integrated in a ubiquitous user model can enrich learner profile from adaptive educational systems. Our framework

Table 1. Information needed to enrich the learner profile

Information category	Related concepts	IMS-LIP
Interests	Interest, product	Interest: information of hobbies and recreational activities
Goals	Purpose of the work, learning goal, information need	Goal: learning, career and other objectives and aspirations
Background	Name, gender, birthdate, place of birth	Identification: biographic and demographic data of a learner or a group
	User's profession, job responsibilities Education and training, military or civic service Work experience	Activity: activities related to learning in any state of completion
	Qualifications, licenses and certifications	Qcl: qualifications, licenses and certifications granted by acknowledged authorities
	Academic transcript	Transcript: summary of academic achievement
	Skills, knowledge and abilities	Competency: skills, abilities and knowledge, acquired in the cognitive, affective and/or psychomotor domains. User´s areas of experience, work history, formal and informal training
Individual traits	Cognitive and learning styles	Accessibility: language proficiency and preference, disabilities, eligibility preferences

described in Section 2.4 can be a mediator between heterogeneous profile providers and adaptive educational system as profile consumers. The process of concept alignment builds bridges between the ubiquitous user model and new consumers finding semantic mappings that enable interoperability. The best suited values gathered from frequently used systems and applications like social networks, FOAF, personal health record, personal devices and educational systems can be delivered on demand to enrich learner profiles.

When sharing and reusing user profile information to enrich learner profile, we are interested in minimizing the cost and effort to build a model, and improving the quality of it.

3.2 Information Needed

As we briefly described in Section 3, educational systems model different aspects of the learner with the aim of personalization. In this paper we are focusing in the learner profile, so we are only addressing the user model. The domain model and the learner's knowledge of the domain are not taken into account. If we want to enrich the learner profile, it is necessary to establish what information is needed and useful for adaptive educational systems.

Brusilovsky *et al.* [2] determined that the five most popular and useful features viewing the user as individuals are: the user's knowledge, interests, goals, background and individual traits. It is important to also take into account learner profile standards like IMS LIP and IEEE PAPI, to make sure that we are considering information in accordance to these two important specifications. The authors of [46] build exact and partial match mappings between PAPI and LIP. Chatti *et al.* presented an investigation of existing learning profile standards and determined that IMS LIP specification is quite complete. IEEE PAPI, for example, does not consider goals and interests as important learner features. For our application scenario, we related the type of information determined relevant in [2] excluding knowledge with LIP structure, and provide examples of related concepts in Table 1. Other categories of information worth considering are context of work, relationships, passwords and security codes.

Table 2. Sources of information to enrich learner profile

Information Category	Sources	Concept Examples
Interests	Social Network Applications FOAF IMS-LIP Adaptive Hypermedia Systems	fb:interests foaf:topic_interest lip:interest
Goals	IMS-LIP Adaptive Educational Systems Intelligent Tutorial Systems	lip:goals
Background	Social Network Applications (LinkedIn, Facebook, Google+) FOAF HR-XML IMS-LIP IEEE-PAPI	linkedIn:position, linkedIn:education foaf:age,foaf:gender, foaf:publications hr-XML:positionHistory lip:activity papi:performance
Individual traits	Social Network Applications (LinkedIn, Facebook, Google+) Personal Health Record Adaptive Educational Systems Intelligent Tutorial Systems	linkedIn:language,fb:language, google+:language MSVault:Condition lip:accessibility papi:preference

Personal context, devices used and user location, affective state and social context determine when, how and with whom the user is confronting a learning task. These concepts are related with LIP's Affiliation, Securitykeys and Relationships specifications.

Table 1 defines the categories of information considered useful in a learner profile given its relation with experts' opinions and standard specifications.

3.3 Heterogeneous Sources of Learner Profile

Once we determined the information needed when trying to enrich a learner profile, it is important to verify if this type of information is available in frequently used systems and applications like social networks, and commonly adopted specifications. Table 2 shows examples of potential sources that could be profile information providers. The intention of this list is to present examples of valuable information scattered in heterogeneous sources. If this information is gathered in a ubiquitous user model it can be shared and reused by adaptive educational systems for personalization and adaptation purposes.

In order to effectively share and reuse heterogeneous sources information, it is necessary to use the providers' transfer mechanisms of develop them (for example, for IMS LIP), obtain the application security tokens

Fig. 3. Use cases representation diagram

and credentials, and solve syntactic and semantic heterogeneities. The task of mapping different sources concepts is very complex and time-consuming and is frequently done by a human expert. Finding exact or partial mappings is not easy even for the experts due to the different semantic and structural conflict described by [39]. The framework described in Section 2.4 is an on-going project that tries to find automatic mappings for concepts from heterogeneous sources.

4 Demonstration of Learner Profile Enrichment

The learning management system Moodle is used in Universidad Panamericana in Mexico as a core component to create online learning sites and manage courses to leverage learning. Essential learner profile information is populated from the university ERP repository, and the learner optionally completes her/his profile. Unfortunately, of about ten thousand students only 30% of them explicitly input some of the profile information (most of the cases only the student's photograph), and very few input information in at least two or more profile fields. In contrast, many students enrich at least one social network profile frequently.

For our proof of concept demonstration, we propose the experiment of harvesting profile information from social network applications, FOAF, personal health records and profiles elaborated according to IMS_LIP and use this information to enrich Moodle's basic learner profile. For the purpose of representing the use cases, Unified Modelling Language (UML) was adopted. Fig. 3 is the use cases representation diagram.

4.1 Enrichment of Moodle's Learner Profile

For our proof of concept demonstration, profile information from different sources was previously extracted and integrated to the ubiquitous user model of our framework. Basic profiles from three social network applications were considered: Facebook, LinkedIn and Google plus. It was also recollected basic demographic information from Microsoft HealthVault as a personal health record
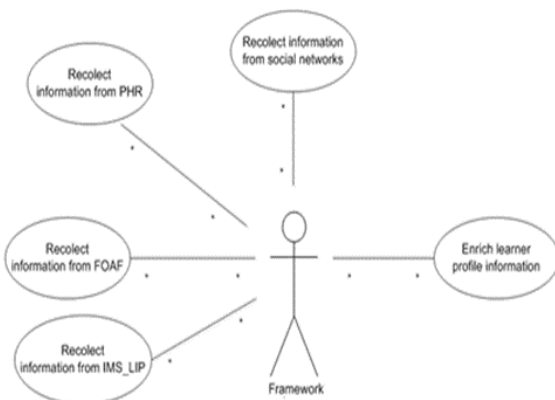
Table 3. Confusion matrix when aligning Moodle's basic learner profile

		Expected Matches	
		Positive	Negative
Outcome of the Process of Concept Alignment	Positive	TP = 9	FP = 2
	Negative	FN = 1	TN = 2

example. Following IMS_LIP specifications, sample instances and the corresponding concept schemas of *accessibility* and *interest* were included in the ubiquitous user model set up. An instance and concept scheme of FOAF was also included in the ubiquitous user model ontology setup. Therefore as a requirement for the enrichment process, the profile information and corresponding concept schemas of the above mentioned suppliers are integrated to the ubiquitous user model ontology.

It is necessary to deal with Moodle's web services in order to enable the extraction of a basic learner profile, and after the performance of the learner profile enrichment, its update in the learning management system. Moodle Rest web services were used with a PHP client. Permissions were explicitly given by the user to obtain authorizations to get access and update all the previously mentioned user profiles.

For the evaluation of the proof of concept demonstration, we focus in the efficiency and effectiveness of the mappings and recommendations resulting of the process of concept alignment of our framework following the metrics proposed in [50]. The human effort required to verify the correctness of the mappings determined by our framework (quantified with the metric *overall*), and the quality of the mappings (quantified with *precision, recall, f-measure* and *fall-out*) will determine how successful was the enrichment of the learner profile. These metrics are based on the notions of true positives (TP), false positives (FP), true negatives (TN) and false negatives (FN).

Regarding the quality of alignment required for this demonstration, it is important that the right exact matches are found. It is a serious error if a wrong exact match is found in the learner profile

matching process, because either the interchange is not possible or unrelated values are populated in a profile learner field. It is a less serious error if no match is found even if the match is relevant. This means that *precision* is expected to be high given that *false positives* are serious mistakes. *Recall*, in turn, is needed to be medium, because it is a less serious error if some expected match is not found (*false negative*).

4.2 Experiment and Results

In the experiment, a corresponding concept scheme was automatically created for the learner profile of Moodle learning management system. The semantic mappings between this consumer's concept scheme and the previously integrated ubiquitous user model were determined by the process of concept alignment in order to enable the interoperability process. In this case, Moodle's learner basic profile includes 14 concepts in addition to the *id* and *username*. No custom fields are taken into account.

The outcomes were evaluated by a human expert who decided if the semantic relations found were correct and recommendations made sense. The evaluation and results of the matching process are shown in Table 3 and Table 4.

Table 3 presents the confusion matrix resulting of aligning Moodle´s basic learner profile and the ubiquitous user model. The exact match was correctly found for six concepts (*firstname, lastname, email, lang, city, country*) and two concepts were identified as close match (*timezone, interest*). When the concepts are considered interchangeable, the match is determined as exact. If the concepts are considered of equivalent meaning, but the similarity of the concept types are not exactly equivalent, the match is defined as close. The concept *mailformat* was found highly related to the collection of concepts in the target, but no exact or close match was found, so the process of concept alignment recommended to add this concept to the collection in the ubiquitous user model (a correct recommendation). This nine alignment outcomes were considered **TP**. Two concepts were discarded correctly given that no match was found for these concepts (*auth, theme*), so it is not feasible to populate them. The

matches for the concepts *auth* and *theme* were correctly not found. This two outcomes were determined **TN**. The labels of the concepts *idnumber* and *description* are used by autonomous user profiles providers and consumers with different meanings. The concept type and structure does not help to disambiguate the meaning in these cases. Erroneous exact matches were found for these concepts, so the outcomes were determined **FP**: The word meaning disambiguation is very difficult to detect even by a human expert. Only when the values of the concepts are interpreted as unrelated, the error is discovered. No match was found for the concept *preferences* even though the ubiquitous user model had this information available. Even though the semantic match is discovered as exact, the internal structure of the provider's concepts is not similar so no match was established. This concept is frequently used with different meaning and complex structures so further transformations are needed in order to enable interchangeability. The outcome of this match was considered **FN**.

The efficiency and effectiveness measuring results for the process of aligning Moodle´s basic learner profile with the ubiquitous user model is shown in Table 4.

Table 4. Efficiency and effectiveness results (Moodle basic profile)

Measure	Metric	Results
Quality of generated mappings (Effectiveness)	Fall-out	50%
	Precision	82%
	Recall	90%
	F-measure	86%
Human effort (Efficiency)	Overall	70%

It was established in Section 4.1 that high *precision* and medium *recall* are needed for the quality of generated mappings. The results of the effectiveness of our experiment show that *recall* result exceeded our requirements and *precision* is acceptable. The fall-out measure calculates the rate of incorrectly discovered matches out of those not expected. Preferably *fall-out* equals 0%, in our experiment *fall-out* result was 50%. The trade-off between *precision* and *recall* is 86% (*F-measure*) when the same weight is given to these two measures.

The *overall* metric quantifies the human effort to correct *false positives* and *false negatives*. Given that the resulting *overall* is 70% (the greater the *overall*, the less effort), the accuracy of the matches can be considered satisfactory for an automatic process.

From these first results it was concluded that the alignment of the learner profile performed well in the proof-of-concept experiment enabling the interoperability between heterogeneous sources. Enrichment of Moodle's basic learner profile sharing and reusing profile information of other domain applications through the mediation of our framework is possible.

5 Conclusions and Future Work

We presented an application scenario of sharing and reusing valuable user profile information from distributed heterogeneous sources to enrich learner profile through the mediation of our ubiquitous user model. Our framework integrates information scattered in social network applications, a personal health record, FOAF and instances of accessibility and interest following IMS_LIP specification. This information was integrated in a ubiquitous user model and used to enrich Moodle basic learner profile. We addressed the syntactic and semantic heterogeneity through the mediation of the ubiquitous user model building semantic mappings with the process of concept alignment to enable interoperability. Although standards like IMS LIP and IEEE PAPI can help establishing specifications as a common structure for interoperability, they are not enough to solve the interoperability problem.

Given the autonomy of the diverse profile providers and consumers in a dynamic distributed multi-application environment, a commonly agreed standard is not feasible even in the same domain. Mediation approaches have to deal with syntactic and semantic heterogeneity, and therefore provide conversion mechanisms between each pair of user models that participates in the interoperability process with the disadvantage of scalability limitation and great

effort if a new stakeholder arrives. A mixed approach was proposed to bridge the gap between mediation and semantic standardization.

The first results show that the effectiveness and efficiency of the process of concept alignment were satisfactory, and the enrichment of the learner profile was enabled. The proposed framework can be a mediator between heterogeneous profile providers and adaptive educational system as profile consumers. Sharing and reusing profile information between user models relieve the learner from repeat configurations, helps deal with the "cold start" problem of new adaptive educational systems. The main benefit of the enrichment of learner information is obtaining a better understanding of the learner with the least intervention and effort the interoperability stakeholders.

New experiments that take into account learner profile standards and various adaptive educational systems must be designed. More empirical evaluation is necessary to refine the model and prove that the solution is practical and generalizable.

References

1. **Brusilovsky, P., Kobsa, A., & Nejdl, W. (2007).** The Adaptive Web. *Methods and Strategies of Web Personalization. Lecture Notes in Computer Science*, 4321. doi: 10.1007/978-3-540-72079-9.

2. **Brusilovsky, P. & Millán, E. (2007).** User models for adaptive hypermedia and adaptive educational systems. *The Adaptive web. Lecture Notes in Computer Science*, 4321, 3–53.

3. **Paulsson, F. (2005).** *Achieving Learner Information interoperability: Position paper for the LIFE expert meeting on Learner Information.* Cambridge, UK.

4. **Brusilovsky, P. (2001).** Adaptive Hypermedia. *User Model User-Adapted Interaction,* 11(1-2), 87–110.

5. **Berkovsky, S., Heckmann, D., & Kuflik, T. (2009).** Addressing Challenges of Ubiquitous User Modeling : Between mediation and semantic integration. *Advances in Ubiquitous User Modelling. Lecture Notes in Computer Science,* 5830, 1–19. doi: 10.1007/978-3-642-05039-8_1.

6. **Carmagnola, F., Cena, F., & Gena, C. (2011).** User model interoperability: a survey. *User Model and User-Adapted Interaction,* 21(3), 285–331. doi: 10.1007/s11257-011-9097-5.

7. **Aroyo, L., Dolog, P., Houben, G.J., Kravcik, M., Naeve, A., Nilsson, M., & Wild, F. (2006).** Interoperability in Personalized Adaptive Learning. *Educational Technology and Society*, 9(2), 4–18.

8. **Bakaev, M., Avdeenko, T. (2013).** Indexing and Comparison of Multi-Dimensional Entities in a Recommender System based on Ontological Approach. *Computación y Sistemas, 17(1),* 5–13.

9. **Carmagnola, F. (2009).** Handling Semantic Heterogeneity in Interoperable Distributed User Models. *Advances in Ubiquitous User Modelling. Lecture Notes in Computer Science,* 5830, 20–36. doi: 10.1007/978-3-642-05039-8_2.

10. **IMS Global Learning Consortium Inc. (2001).** *IMS Learner Information Package Specification.* Retrieved from http://www.imsglobal.org/profiles/.

11. **Ramandalahy, T., Vidal, P., & Broisin, J. (2009).** Opening Learner Profiles across Heterogeneous Applications. *Ninth IEEE International Conference on Advanced Learning Technologies (ICALT 2009),* Riga, Latvia, 504–508.

12. **Martinez-Villaseñor, M. de L., Gonzalez-Mendoza, M., & Hernandez-Gress, N. (2012).** Towards a Ubiquitous User Model for Profile Sharing and Reuse. *Sensors,* 12(10), 13249–13283. doi: 10.3390/s121013249.

13. **Martinez-Villaseñor, M. de L. (2013).** *Design and Implementation of a Framework for Ubiquitous User Model Interoperability.* PhD Thesis, Instituto Tecnológico y de Estudios Superiores de Monterrey, Estado de México, México.

14. **Dolog, P. & Schäfer, M. (2005).** A framework for browsing, manipulating and maintaining interoperable learner profiles. *User Modelling 2005. Lecture Notes in Computer Science*, 3538, 397–401.

15. **Musa, D.L. & De Oliveira, J.P.M. (2005).** Sharing learner information through a web services-based learning architecture. *Journal of Web Engineering*, 4(3), 263–278.

16. **Brusilovsky, P. (2004).** KnowledgeTree : A Distributed Architecture for Adaptive E-Learning. *13th International World Wide Web Conference on Alternate track papers & posters.* New York, USA, 104–113.

17. **Brusilovsky, P., Sosnovsky, S., & Yudelson, M. (2005).** Ontology-based Framework for User Model Interoperability in Distributed Learning Environments. *World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education.* Chesapeake, VA, 2851–2855.

18. **Assad, M., Carmichael, D.J., Kay, J., & Kummerfeld, B. (2007).** PersonisAD: Distributed, active, scrutable model framework for context-aware services. *Pervasive Computing. Lecture Notes in Computer Science*, 4480, 55–72.

19. **Kay, J. (1994).** The um toolkit for cooperative user modelling. *User Model User-adapted Interaction,* 4(3), 149–196.

20. **Kay, J. (1999).** *A scrutable user modelling shell for user-adapted interaction*. PhD Thesis, University of Sydney, Australia.

21. **Kay, J. (2000).** Accretion Representation for Scrutable Student Modelling. *Intelligent Tutoring Systems. Lecture Notes in Computer Science*, 1839, 514–523.

22. **Kay, J., Kummerfeld, B., & Lauder, P. (2002).** Personis: a server for user models. *Adaptive Hypermedia and Adaptive Web-Based Systems. Lecture Notes in Computer Science,* 2347, 203–212.

23. **Distributed Management Task Force. (2003).** CIM User and Security Model White Paper.

24. **McCalla, G., Vassileva, J., Greer, J., & Bull, S. (2000).** Active Learner Modelling*. Intelligent Tutoring Systems. Lecture Notes in Computer Science,* 1839, 53–62.

25. **Vassileva, J. (2001).** Distributed user modelling for universal information access. *9th International Conference on Human-Computer Interaction.* New Orleans, USA, 122–126.

26. **Niu, X., Mccalla, G., Vassileva, J. (2003).** Purpose-based User Modelling in a Multi-agent Portfolio Management System. *User Modeling 2003. Lecture Notes in Computer Science*, 2702, 398–402.

27. **Vassileva, J., Mccalla, G., Greer, J. (2003).** Multi-agent multi-user modeling in I-Help. *User Model and User-adapted Interaction*, 13(1-2), 179–210.

28. **Stewart, C., Cristea, A., Celik, I., Ashman, H. (2006).** Interoperability between AEH user models. *Proceedings of the joint international workshop on Adaptivity, personalization & the semantic web*. Odense, Denmark, 21–30.

29. **Trella, M., Conejo, R., Guzman, E., & Bueno, D. (2003).** An Educational Component Based Framework for Web ITS Development*. Web Engineering. Lecture Notes in Computer Science, 2722*, 134–143.

30. **Van der Sluijs, K. & Houben, G.J. (2006).** A generic component for exchanging user models between web-based systems. *Learn International Journal of Continuing Engineering Education and Lifelong Learning,* 16(1-2), 64–76.

31. **González, G., López, B., & de la Rosa, J.L. (2005).** A Multi-agent Smart User Model for Cross-domain Recommender Systems. *Workshop beyond Personalization IUI'05*, San Diego, Ca.

32. **Pazzani, M.J. (1999).** A Framework for Collaborative, Content-Based and Demographic Filtering. *Artificial Intelligence Review,* 13(5-6), 393–408. doi: 10.1023/A:1006544522159.

33. **Berkovsky, S. (2006).** Decentralized Mediation of User Models for a Better Personalization. *Adaptive Hypermedia and Adaptive Web-Based Systems. Lecture Notes in Computer Science*, 4018, 404–408.

34. **Berkovsky, S., Kuflik, T., Ricci, F. (2008).** Mediation of user models for enhanced personalization in recommender systems. *User Modeling and User-adapted Interaction,* 18(3), 245–286. doi: 10.1007/s11257-007-9042-9.

35. **Berkovsky, S., Kuflik, T., & Ricci, F. (2007).** Cross-domain mediation in collaborative filtering. *User Modeling 2007. Lecture Notes in Computer Science,* 4511, 355–359.

36. **Berkovsky, S., Kuflik, T., & Ricci, F. (2009).** Cross-representation mediation of user models. *User Model User-adapted Interaction,* 19(1-2), 35–63. doi: 10.1007/s11257-008-9055-z.

37. **Berkovsky S, Gorfinkel A, Kuflik T, Manevitz L (2006)** Case-based to content-based user model mediation. Work. Ubiquitous User Model. Riva de Garda, Italy, pp 1–4.

38. **Sosnovsky, S., Brusilovsky, P., Yudelson, M., Mitrovic, A., Mathews, M., & Kumar, A. (2009).** Semantic Integration of Adaptive Educational Systems. *Advances in Ubiquitous User Modelling. Lecture Notes in Computer Science,* 5830, 134–158.

39. **Bouzeghoub, A. & Gancarski, A.L. (2011).** An Agent-Based Service Architecture for User Profiles Dynamic Share. *Seventh International Conference on Semantics Knowledge and Grid (SKG),* Beijing, China, 137–143.

40. **Walsh, E., O'Connor, A., & Wade, V. (2011).** Supporting Learner Model Exchange in Educational Web Systems. 7th *International Conference on Web Information Systems and Technologies,* Noordwijkerhout, The Netherlands.

41. **Shvaiko, P., Giunchiglia, F., Schorlemmer, M., McNeill, F., Bundy, A., Marchese, M., Yatskevich, M., Zaihrayeu, I., Ho, B., Lopez, V., Sabou, M., Abian, J., Siebes, R., & Kotoulas, S. (2006).** *Openknowledge Deliverable 3.1.:*

Dynamic Ontology Matching: A Survey (DIT-06-046). Trento, Italy: University of Trento.

42. **Miles, A. & Pérez-Agüera, J.R. (2007).** SKOS: Simple Knowledge Organization for the Web. *Cataloging & Classification Quarterly,* 43, (3-4), 69–83.

43. **Dice, L.R. (1945).** Measures of the amount of ecologic association between species. *Ecology,* 26(3), 297–302.

44. **Fellbaum, C. (2005).** WordNet and Wordnets. In: Brown, K. E (ed). *Encyclopedia of Language and Linguistics* (665–670), Oxford, UK: Elsevier.

45. **Kravčík, M. & Gašević, D. (2007).** Leveraging the Semantic Web for Adaptive Education. *Journal of Interactive Media in Education*, Special Issue: Adaptation and IMS Learning Design.

46. **Chatti, M.A., Klamma, R., Quix, C., & Kensche, D. (2005).** LM-DTM: An Environment for XML-Based, LIP/PAPI-Compliant Deployment, Transformation and Matching of Learner Models. *Fifth IEEE International Conference on Advanced Learning Technologies (ICALT 2005). Kaohsiung*, Taiwan, 567–569.

47. **The Nielsen Company. (2012).** August 2012: Top U.S. Web Brands and Technology News Websites. Retrieved from http://blog.nielsen.com/nielsenwire/online_mobile/august-2012-top-u-s-web-brands-and-technology-news-websites/.

48. **Russell, M.A. (2011).** *Mining the Social Web: Analyzing Data from Facebook, Twitter, LinkedIn and Other Social Media Sites.* Sebastopol, CA: O'Reilly.

49. **Ding, L., Zhou, L., Finin, T., & Joshi, A. (2005).** How the Semantic Web is Being Used : An Analysis of FOAF Documents. *38th Annual Hawaii International Conference on System Sciences*. Big Island, HI, USA, 4, 113.3.

50. **Bellahsene, Z., Bonifati, A., Duchateau, F., & Velegrakis, Y. (2011).** On Evaluating Schema Matching and Mapping. *Schema Matching and Mapping* (253–291), New York: Springer.

Ma. Lourdes Martínez-Villaseñor received her Ph.D. in Computer Science from *Tecnológico de Monterrey, Campus Estado de México*. She is Researcher-Professor and Head of Computer Science Department at *Universidad Panamericana*, Mexico.

Miguel González-Mendoza received his Ph.D. in Computer Science from *Institut National des Sciences Appliquées* de Toulouse. He is Head of the Graduate Programs in Computer Sciences at *Tecnológico de Monterrey*, Campus Estado de México.

Ignacio Danvila Del Valle received his Ph.D. in Business and Economic Science from *Universidad Complutense de Madrid*. He is Researcher-Professor at *Universidad Complutense de Madrid*.