Adult Student Modeling for Intelligent Distance Learning Systems
Panagiotopoulos Ioannis, Kalou Aikaterini, Pierrakeas Christos and Kameas Achilles

Educational Content, Methodology and Technology Laboratory (e-CoMeT Lab)
Hellenic Open University, Patras, Greece
{gpanagiotopoulos, kalou, pierrakeas, kameas}@eap.gr

Abstract: One of the most important components in a learning support system is the learner model, as it contains useful information about an individual such as learning preferences and academic performance. The goal of the research presented in this paper is to define how a learner model can be distributed with the help of semantic web technologies, based on stereotypes as a useful mechanism for the initialization of an intelligent learning system. These stereotypes have been derived from an empirical study on a sample of adult learners at a distance learning University, while the proposed model also reflects features from several standards for a learner modeling. Finally, a web application is presented, in order to evaluate the learner model and test the automatic categorization of learners into stereotypes according to their basic characteristics.

Keywords: Ontology, Intelligent Tutoring Systems, Stereotypes, Personalized Learning, Learner Model

1. Introduction

Intelligent Tutoring Systems (ITSs) are complex systems that can be adapted easily to each student’s cognitive features, characteristics and learning progress [1, 31]. These systems use a large amount of educational knowledge and many of them also employ pedagogical methodologies.

Traditional ITSs consist of the following four modules: (a) the domain module, which contains all the knowledge (educational content), (b) the student model, where learner’s specificities and progress are stored into, (c) the pedagogical module, which contains all the information relevant to the various pedagogical decisions and (d) the user interface
which enables communication between the user and the system [34]. Especially, in the case of multi-agent architectures, the interaction between these modules is achieved through the communication of intelligent agents assigned to each module. For example, a learner model agent is responsible for answering queries from other agents about learner’s information, which information is included in the student model.

Since, ITSs are systems par excellence that offer student-centered learning, the student model constitutes the core module of these systems. The characteristics and progress of the students are captured in the student model. This is achieved by using AI techniques to represent pedagogical decisions, domain knowledge, and personal information about the student [31]. Since there are many candidate characteristics of a user that can be included in the student model, the selection of the appropriate set of characteristics is a very challenging and significant procedure. Consequently, we have to obtain a tradeoff of the completeness of the model so that the systems can be adapted successfully and their performance is not affected. Some of the basic student characteristics maintained in the model are: (a) demographic information, (b) knowledge of the teaching domain, (c) background and interests, (d) learning styles and interaction preferences as well as (e) learning goals and specificities that can affect the learning procedure. From the aforementioned characteristics (a) is a basic feature when describing a user, while (b) and (c) are essential parameters in every educational process. Characteristics (d) and (e) are considered as crucial for a user-centered intelligent tutoring system in order to deliver the appropriate educational material, according to the individual’s needs.

In the context of our work, we made an attempt to propose a student model that could be utilized in the ITS that is going to be developed for the needs of the Hellenic Open University. Since, the Hellenic Open University is a foundation, which is specialized in the life-long and distance learning; we took into account the special features of adult students for the design of the proposed student model. In order to identify more efficiently the learner characteristics, an empirical study with the contribution of an expert coming from the science of Anthropology was conducted on a sample of students
of the Hellenic Open University. Moreover, we represent all the elements of the student model using the notion of ontologies.

Briefly, the proposed model:

• is highly adapted to the particular characteristics of the student, such as the learning style,
• is adult student-oriented as it includes characteristics like previous experience and learning goals,
• is not designed as a model for collecting data and serve as a classic CV
• is based on an empirical study by a social scientist among adult students
• combines different modeling approaches for representing the information about the student

The rest of the paper is structured as it follows. In Section 2, we discuss the student modeling approaches based on ontologies and the student modeling standards. In Section 3, we elaborate on the proposed student model, in terms of the learning styles, the modeling approaches and the basic characteristics of the student. Next, we briefly discuss the main differences of the proposed model with the existing standards. In Section 4, we outline the ontology that represents the proposed student model. In Section 5, we present the ontology-based application that has been implemented in order to populate the proposed student model ontology via a simple user interface and finally, in Section 6, we conclude with directions for future work.

2. Related Work

Student modeling constitutes a promising research topic that has matured over the years and ensures characteristics such as personalization and adaptivity in the context of e-learning systems. Recently, student modeling researchers have concentrated on the adoption of semantic web technologies and principles, since ontologies (pillar of Semantic Web) are becoming the standard way of knowledge representation on Web.
The concept of Ontology initially was derived from philosophy, where it means the study of being or existence. Several definitions of the formal ontology exist. The most cited definition was provided in [15] and later modified by Guarino in [16]: “Ontology is an explicit intentional specification of shared conceptualization of the domain”. In knowledge-based systems an ontology is a set of representational terms that associate the names of entities in the universe of discourse (classes, relations, etc.) with human-readable text, describing what the names are meant to denote [14].

Ontologies have been widely used for student modeling for two primary reasons: (a) ontologies support the formal representation of abstract concepts and properties in a way that they can be reused by many tasks or extended if needed and (b) they enable the extraction of new knowledge by applying inference mechanisms (e.g. reasoner) on the information presented in the ontology. Winter et al. [36] propose a set of best practices for ontology-based student modeling and summarize the advantages of ontology-based models in “formal semantics, easy reuse, easy portability, availability of effective design tools, and automatic serialization into a format compatible with popular logical inferences engines”. Moreover, Peña and Sossa [30] pretend that ontologies facilitate the reuse and the integration of resources and services, so that e-learning systems can provide better applications.

Therefore, a plethora of ontology-based approaches for student modeling have been proposed in the field of ITSs. Paneva [29] proposes an ontology-based student model for eLearning systems that adopts technologies and standards from the Semantic Web. Chen et al. [3] describe a domain-independent student model for a multi-agent intelligent educational system (IES). In [32], they propose a student model ontology for an e-Learning system. The ontology is based on the representation of prior knowledge of the student and his/her learning style. In [21] they propose an ontology-based learning environment, called TANGRA. The representation of the educational material is based on learning objects (LOs), while the student model, the system uses, is based, on popular student modeling standards. Jeremic et al. [19] describe a student model for the Design Pattern intelligent tutoring system. Similarly, many other approaches have been presented...
in literature (i.e. [18], [25]). Additionally, in [5] the authors propose a student modeling mechanism for Intelligent Virtual Environments for Training (IVETs). They divide the student information in three major categories: (a) student profile (personal data), (b) state of student's progress and (c) trace of student's activity. The Personal Reader [8] exploits the strengths of Semantic Web so as to represent information about students, which is required to recommend appropriate learning resources taking into account student interests/preferences, student performance in different courses and student goals. SoNITS (Social Network for Information Technology Students) [26], incorporates an ontology-based student model which concerns Information Technology (IT) field and includes only industrial IT skills. The OPAL project (Ontology-based framework for personalized adaptive learning) [4] combines ontologies and learner models for a Web-based Java programming course. The student model includes preferences set by the student himself as well as learning progress and performance.

Furthermore, apart from the ontology-based student models that we described above, many attempts have been made in order to model the learner data in a more formal way and have been resulted in a number of standards, such as PAPI (Public and Private Information) [23], IMS LIP (Learner Information Package) [35], eduPerson [Error! Reference source not found.], Dolog LP [9, 7], FOAF (Friend of a Friend) [2] etc. Even if these models share a set of common learner characteristics, they vary on their main purpose and the way in which a system may use their embedded information. It is a usual practice to produce a learner profile for a learner system combining different learner standards and profiting from their unique benefits.

The PAPI Learner Standard describes a particular subset of all possible types of learner information. Learner information is considered a subset of general information about learning technology. In the last edition of the PAPI Learner Standard there were six information types. The category Personal contains information about names, contacts and addresses of a learner. Relations category is about the learner's relationship to other users of learning technology systems, such as teachers, proctors, and other learners. Security aims to provide slots for the learner's security credentials, such as: passwords,
challenge/responses, private keys, public keys, biometrics, etc. Portfolio is for representing learner’s previous works or references. Performance category is for storing information about learner’s performance through the educational process. Finally, Preference is for describing learner’s preferences that may improve human-computer interactions. Figure 1 depicts the conceptual view of the PAPI standard.

![Figure 1. The main categories of the PAPI standard.](http://wwwis.win.tue.nl/ah2003/proceedings/um-1/)

Similarly the IMS LIP standard contains several categories for data about a user. The identification category provides information about learner’s biographic and demographic data relevant to learning. Goal category represents learner’s career and learning objectives and aspirations. The QCL category is used for the identification of qualifications, certifications and licenses granted by recognized authorities. The activity category may contain any learning-related activity in any state of completion. This category includes formal and informal education, training, work experience, and military or civil service. The interest category can contain any information describing hobbies and recreational activities. The competency category serves as a slot for skills, knowledge, and abilities acquired in the cognitive, affective, and/or psychomotor domains. Affiliation category represents information about membership in professional organizations. The
The accessibility category aims for general accessibility to the learner information as defined through language capabilities, disabilities, eligibilities and learning preferences. The transcript category represents an institutionally-based summary of academic achievement. Finally Security key is for setting the passwords and security keys assigned to the learner for transactions with learner information systems and services. All the above categories are depicted in figure 2.

Figure 2. The core categories for learner data in IMS LIP. (IMS Learner Information Packaging Information Model Specification. Retrieved November-12-2012, from: http://www.imsglobal.org/profiles/lipinfo01.html)

Dolog LP is a learner profile which was proposed by Dolog and Nejdl that uses ontologies to semantically enhance learning systems in order to provide personalized services. The profile takes advantage of the semantic tools RDF\(^1\) (Resource Description Framework) and RDF Schema\(^2\) in order to model learners and to be extendable. The proposed learner model is based on the two standards PAPI and IMS LIP and includes five main categories to describe learner’s characteristics: (a) Other User Features

---

1. [http://www.w3.org/RDF/](http://www.w3.org/RDF/)
2. [http://www.w3.org/TR/rdf-schema/](http://www.w3.org/TR/rdf-schema/)
describes student’s Preferences (language, proficiency, etc.), Goals and Interests; (b) Study Performance holds information about learner’s Performance, Portfolio and Certification; (c) Identification provides information about the learner (name, email, address and telephone); (d) Human Resource Planning (HRP) holds information about the organization in which the learner is a member of; (e) Calendar includes information about learner’s appointments and events.

eduPerson is a specification released by Internet2\(^3\) and EDUCASE\(^4\). This schema was designed to include widely-used personal and organizational attributes and thus facilitate the communication between higher education institutes. The information covered by this standard is mainly classified in the following categories: (a) General attributes, which cover general information about the learner, such as name, home address, phone number, security information, and information about the organization in which the learner is a member of; (b) attributes such as person’s affiliation, person’s ID, and generally information that facilitate collaboration between institutions.

FOAF is a RDF vocabulary for describing people, documents and organizations and it was mainly developed in order to build communities and social groupings [2]. FOAF vocabulary distinguishes five basic categories (classes) for describing a person: (a) Person describes basic information about a person such as name, email, age, etc.; (b) Document and Image holds information about a document or an image; (c) Organization holds information about the social institutions the person is a member of; (d) OnlineAccount provides information about the accounts a person has; (e) Projects and Groups holds information about the projects or groups a person is a member of.

3. Student Model Development

In this section, we shall give a brief description of the basic components of the proposed approach and some of the basic characteristics of the model. We also set out some of the most commonly used approaches for representing a student model.

\(^3\) http://www.internet2.edu/
\(^4\) http://www.educause.edu/
3.1 Learning style

One of the most important components of the student model in an ITS is the personal learning style of the learner. The term “learning style” is used to describe the individual differences in the learning process. More precisely, with the available learning styles we can indicate all the different ways that students can learn. It is based on the assumption that each person has a unique and distinctive way to learn, i.e. to collect, process and organize information [20].

Among the models and theories presented in the literature, we have adopted the Felder-Silverman theory for student modeling. Most existing learning-style based theories classify students into few coarse grained groups, whereas Felder and Silverman describe the learning styles of a student in more detail, distinguishing between preferences on four dimensions [11].

![Figure 3. The Felder-Silverman model. (Melissa Cater, 2011. Incorporating Learning Styles into Program Design, Retrieved November-12-2012, from: https://lsuagcenterode.wordpress.com/2011/08/16/incorporating-learning-styles-into-program-design/)](image-url)
According to the Felder-Silverman model, the learning types are categorized in the following four dimensions: (a) active/reflective, (b) sensing/intuitive, (c) visual/verbal, and (d) sequential/global. In figure 3 we can see the complete model along with the descriptions for each dimension.

3.2 Modeling Approach

There is a variety of techniques and classic approaches to represent the acquired learner information. The most common representation of a student model is the overlay model. The overlay model represents a learner's knowledge as a subset of the domain knowledge (expert's knowledge). Therefore, the system provides the learner with educational material until learner's knowledge coincides with the expert's knowledge [6]. Another approach which is widely used is the buggy model. Systems that use such models record and represent the most common/frequent mistakes made by learners based on statistics. Finally, one widely adopted approach for student modeling is the use of stereotypes [22, 33]. New learners are classified into distinct categories and the system adjusts its performance based on the category assigned to the learner. In figure 4 are depicted the overlay and buggy models.

![Diagram of overlay and buggy models](http://www.ent.mrt.ac.lk/~ekulasek/en577/Applications%20of%20AI%20in%20Education.htm)

**Figure 4.** Overlay and buggy models. (Beck, J., Ster, M., Haugsjaa, E. *Applications of AI in Education*. Retrieved November-12-2012, from: http://www.ent.mrt.ac.lk/~ekulasek/en577/Applications%20of%20AI%20in%20Education.htm)
In the context of our work, we adopt a combination of the stereotype and overlay techniques. A fully stereotype-based model was excluded as a choice because (1) the initialization of the system derived from students descriptions or questionnaires may not be accurate for every knowledge domain and (2) the system would adapt to the learner’s needs very slowly. So we developed a model where some attributes of the student profile (e.g. previous knowledge, experience in a specific knowledge domain) are initialized based on a stereotype. In addition, dynamic attributes related to the learning process are represented with an overlay model. After the initialization phase, the profile is dynamically modified, as the overlay model is updated with the information gathered by the interaction between learner and system.

3.3 Basic Characteristics of the Students

The users’ classification in categories, called stereotypes constitutes a technique that has been widely used in user modeling systems. Stereotypes can be specified according to the following criteria: age, gender, educational level, working experience etc.

An empirical study was conducted by the Educational Content, Methodology and Technology Laboratory⁵ among students of the Hellenic Open University⁶ (HOU) in order to extract the basic characteristics and formulate the corresponding stereotypes of the students. The HOU was founded in 1992 and its mission is to provide distance learning education at both undergraduate and postgraduate level. Today, HOU has 15,074 undergraduate, 11,243 postgraduate and 68 PhD students. It offers 32 courses and 203 course modules. Additionally occupies 1600 tutors and 42 professors as academic stuff. The basic educational unit of the HOU is called “course module” and covers a specific subject on an undergraduate or postgraduate level. Each course module is equivalent to 3 semestral academic lessons.

---

⁵ http://eeem.eap.gr
⁶ http://www.eap.gr/index_en.php
### Table 1. Dimensions of the student profile and their corresponding values

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning style</td>
<td>active/reflective – sensing/intuitive – visual/verbal – sequential/global</td>
</tr>
<tr>
<td>Use of technology</td>
<td>adaptable - adaptive</td>
</tr>
<tr>
<td>Computer literacy</td>
<td>novice – beginner – advanced</td>
</tr>
<tr>
<td>Previous experience</td>
<td>novice – beginner – advanced</td>
</tr>
<tr>
<td>Time for study</td>
<td>no time – little – much</td>
</tr>
<tr>
<td>Reasons for education</td>
<td>career development – career change – general knowledge</td>
</tr>
<tr>
<td>Academic literacy</td>
<td>poor – good – excellent</td>
</tr>
<tr>
<td>Socialization style</td>
<td>lonely – collaborative</td>
</tr>
</tbody>
</table>

The students who participated in the study were chosen based on their different characteristics such as different gender, age, educational background and current course. The study included (a) personal interviews with the students and (b) observation of the face to face meetings, by a social scientist. In particular, from the 13 students who were interviewed, 5 were male and 8 female, 8 of them pursued undergraduate studies and 5 postgraduate, 10 of them were working in the public sector, 2 were unemployed and 1 was working as a freelancer. Table 1 summarizes the dimensions of the student profile and the corresponding values that came up from the empirical study. The values of each dimension reflect a unique stereotype, e.g. for the dimension “Reasons for education” we use three stereotypes (career development, career change and general knowledge). The classification of students into stereotypes is implemented by SWRL rules in section 4.

Besides the modeling approach that defines the specialization of the model, a few more model characteristics have been taken into account:

- it is a dynamic model that can change over time as the system collects information about the individual and creates a dynamic model for each individual,
• it is a long-term model that keeps generalized information regarding the user-system interactions. In contrast, a short-term model keeps information only for the current interaction of the user with the system,

• it is a combination of “active” and “passive” user model, i.e. in the beginning the user provides directly information about him/her and then the system indirectly collects more information.

Regarding the specialization, the model can represent each individual learner or groups of learners, according to their stereotypic profile.

3.4 Limitations of existing Standards and Models

The proposed ontology-based student model is partially based on the standards that we mentioned in section 2. It is a common belief that PAPI and LIP are the most significant and important among the known standards due to their extended use and the benefits that they provide when used jointly. In [28], the presentation of the main characteristics of the aforementioned standards and the comparison of them denote the importance and the completeness of PAPI and LIP. However, these standards reflect different perspectives on the attributes of a learner.

Table 2 shows a comparison between the learner models we described in Section 2. The comparison is based on the existence of the most important learner attributes that resulted from our research in previous subsections.

From the above brief discussion we can summarize as follows:

• PAPI and IMS LIP focus on the performance and achievements of the learner. Nevertheless, both standards have some shortcomings. For example, the IMS LIP standard is based on the notion of a classic CV, while the PAPI standard considers student's performance as the most important information. In more details, PAPI does not consider the learner’s goals and reasons for education, which can be used for the system’s personalization. However in the context of our work, we took
them into consideration and incorporated some of their basic notions to our proposed student model so that it conforms to these international standards.

- eduPerson is best for collecting data about the learner and communicate this data between institutions.
- FOAF was not initially designed for collecting data. It was designed to describe relations between learners (through the relation “knows”). It can be used for automatic personalization.

Table 2. Comparison of the learner models. The “✓” symbol is used for support of the corresponding attribute

<table>
<thead>
<tr>
<th>Attribute</th>
<th>PAPI</th>
<th>IMS LIP</th>
<th>Dolog LP</th>
<th>FOAF</th>
<th>eduPerson</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal Information</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Learning style</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Computer literacy</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Competency</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time for study</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goals</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Academic literacy</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Socialization style</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Dolog LP seems the model that covers most of the learner’s characteristics, as they have resulted from our research. However it doesn’t cover the category time for study, which is an important attribute for adult learners in distance learning systems. Moreover, the category Calendar doesn’t seem to fit our needs.

Regarding the ontology-based approaches we mentioned in section 2, our idea adopts the basic principles of the student model described in [5] suitably adapted to the needs and characteristics of an adult learner (e.g. time for study, previous experience, educational level and learning goals) in a distance learning educational framework. However, we believe that the rest of the proposed models lack of some important learner’s
characteristics, according to our research. For example, in [29] the author does not include student's learning style in their model, while in [18] only student's performance and his/her interaction with the system is taken into consideration. Furthermore, in [32] the authors do not include student's preferences, learning goals and motivation state in their model.

3.5 Ontology-based Student Model Requirements

The proposed ontology-based model describes the stable information of a learner (e.g. his demographic data), his performance and the degree of knowledge acquisition during a course or course module. In addition, learner’s characteristics that may change dynamically, such as the learning style, interaction preferences and the learning goals are presented. Below there are some competency questions that the student model ontology is able to answer:

**Demographics**

- What is the name of a learner?
- Is a learner a male or female?
- What is the education status of a learner?
- How motivated is a learner?
- What experience does a learner have in a certain topic?
- What's the learning style of a learner?

**Interaction with the system**

- What knowledge has a learner mastered? / Has a learner mastered a certain topic?
- What learning objects were given to a learner?
- What media does a learner use for interaction?
- How is the overall performance of a learner?
- How well does a learner master a certain topic?
As we mentioned in section 3, the proposed model integrates stereotypic profiles of a learner. On the other hand, this model is implemented through ontologies. The contribution of ontologies to the implementation of stereotype-based models can be multiple [12]:

- domain ontologies can be used for the population of the stereotypic profiles
- a single stereotypic profile can be represented as an ontology
- ontologies can contribute to the organization of the stereotypic structure and improve the reasoning among stereotypes.

4. Student Model Development

In this section, we thoroughly describe the *Student Model* ontology that has been developed in order to capture the main concepts presented in section 3. The focus of our attempt is not restricted on modeling the static profile of the user, but encompasses both permanent and dynamic characteristics. Moreover, the developed ontology complies partly with well-known standards for student modeling, i.e. IEEE PAPI Learner [23] and IMS Learner Information Package (LIP) [35].

In order to build the ontology, we followed a widely-adopted methodology, proposed in [27]. As far as its formal representation is concerned, we adopted the Web Ontology Language (OWL) [24], which is a W3C standard. More specifically, our ontology falls into the OWL DL sublanguage, which provides the maximum expressiveness, while maintains computational completeness (all the conclusions are measurable and all calculations are terminated in finite time). Moreover, ontologies based on description logic, are a good choice for the description of information that feeds learning system applications.

The development process of the ontology was accomplished with the aid of Protégé\(^7\) tool. Protégé is an ontology development tool which is the most widely used of all available tools. Its advantages over other tools are: (a) advanced user interface, (b) extensibility

\(^7\) [http://protege.stanford.edu/](http://protege.stanford.edu/)
through plug-ins and (c) wide functionality provided either by the use of plug-ins or by a wide range of different formats that can be exported or imported.

In the proposed ontology Student Model, we define a set of four upper level classes, namely Student, StudentCourseInformation, StudentCurrentActivity and StudentPersonalInformation. The class hierarchy of the ontology, as displayed in Protégé, is depicted in figure 5. The class Student represents any student. The StudentCourseInformation class comprises information relevant to the student’s performance during the overall educational process and has a number of subclasses that are listed below, together with a brief explanation:

A. Assignment - the written assignments that the student has to submit during a course module
B. CourseModule - the course modules of the course program
C. FaceToFaceMeeting - face to face meetings during a course module
D. LearningObject - the learning objects that the student has been taught
E. LearningOutcome - the learning outcomes succeeded by the student as indicated by the learning objects
F. School - the school for which the student is registered
G. WrittenExams – the written exams that the student has to participate during a course module.

In order to capture any detail in terms of a student’s activity for the current academic year, we define the class StudentCurrentActivity. Student’s activity for the current academic year can be specified by the following three axes: (i) current chosen course modules (class CurrentCourseModule), (ii) the experience on a specific course module that the student has previously gained (class PreviousExperience) and (iii) the student’s goals on a specific course module (class SessionGoals).
The more compact class in the proposed ontology, `StudentPersonalInformation`, is defined so as to represent mostly static and permanent student information, describing not only simple data, like demographic data, but more complex characteristics that concern student’s interaction with the e-learning system. Table 3 lists the subclasses that exist under the upper level class `StudentPersonalInformation`. The table gives also a brief description of the entities that are represented by these classes.
<table>
<thead>
<tr>
<th>Class Name</th>
<th>Class Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accessibility</td>
<td>The overall set of features that characterizes the student’s behavior during his interaction with the e-learning system</td>
</tr>
<tr>
<td>Disabilities</td>
<td>The set of student’s disabilities that could affect the educational process</td>
</tr>
<tr>
<td>DemographicData</td>
<td>Student’s demographic data</td>
</tr>
<tr>
<td>InteractionPreferences</td>
<td>Student’s preferences regarding interaction with the e-learning system</td>
</tr>
<tr>
<td>MediaPresentation</td>
<td>Student’s preferences regarding the presentation of learning objects</td>
</tr>
<tr>
<td>Language</td>
<td>Student’s preferences regarding the language of the learning objects</td>
</tr>
<tr>
<td>LanguageSpoken</td>
<td>Student’s native languages</td>
</tr>
<tr>
<td>LanguagePreferred</td>
<td>Language that the student prefers for the presentation of learning objects</td>
</tr>
<tr>
<td>Aesthetics</td>
<td>Aesthetic factors such as the use of highly interactive sensory and visual communication</td>
</tr>
<tr>
<td>Color</td>
<td>Student’s preferences regarding the coloring scheme of learning system’s environment</td>
</tr>
<tr>
<td>Fonts</td>
<td>Student’s preferences regarding the fonts used by the learning system’s environment</td>
</tr>
<tr>
<td>LearningStyle</td>
<td>Student’s learning style - This class will be further divided to the sub classes according to the Felder-Silverman theory</td>
</tr>
<tr>
<td>MotivationState</td>
<td>Student’s motivation during the educational process</td>
</tr>
<tr>
<td>LearningGoals</td>
<td>Overall goals set by the student</td>
</tr>
<tr>
<td>ReasonsForEducation</td>
<td>The reasons why the student desires to engage in the educational process</td>
</tr>
<tr>
<td>AcademicLiteracy</td>
<td>Student’s previous formal educational experiences</td>
</tr>
</tbody>
</table>
Interests  
Student’s interests

TimeStudy  
The average time per day that the student can use for studying

On the other hand, relationships between instances (members of classes) are modeled as object properties. In this context we define a set of object properties (mostly of the hasA kind). This kind of object property is used for expressing the association of the aforementioned characteristics with students. It links an instance of the class Student to instances of classes that reflects student characteristics such as StudentCourseInformation, StudentCurrentActivity, PreviousExperience, DemographicData, StudentPersonalInformation, LearningStyle, InteractionPreferences, Disabilities, Interests, TimeStudy, ReasonsForEducation and AcademicLiteracy. Furthermore, datatype properties, that link individuals to data values, have been set in order to define more effectively the classes. In the proposed model the stereotypic profiles (table 1) have been expressed as datatype properties. In figure 6 are depicted all the datatype properties of the Student Model ontology as shown in Protégé tool.

![Data properties](image)

**Figure 6.** The Student Model ontology as displayed in Protégé
In addition, as foresaid, we have adopted a scheme inspired by the Felder-Silverman Learning Style Model [13], in order to infer the student’s learning style. The eight proposed learning styles are captured as individuals of the class LearningStyle, in the Student Model ontology (see Section 3.1 for a learning styles description).

The proposed ontology has been enriched with a set of rules in order to enable inference mechanisms (i.e. reasoner) to automatically classify the students into different stereotypic profiles (table 1). As foresaid, these rules and the stereotypic profiles have resulted from an empirical research on a sample of adult learners. All the rules are expressed in Semantic Web Rule Language\(^8\) (SWRL) and a subset is given in table 4. These rules respectively indicate:

- **Rule #1**: “if the student is female, over 50 years old and doesn’t have a bachelor degree, then has little familiarity with computers”
- **Rule #2**: “if the student is female, over 50 years old and doesn’t have a bachelor degree, then has much time for study”
- **Rule #3**: “if the student is female, over 50 years old and doesn’t have a bachelor degree, then she is enrolled in the educational process for the acquisition of general knowledge”
- **Rule #4**: “if the student is male, over 30 years old and doesn’t have a bachelor degree, then he is enrolled in the educational process for career development”.

<table>
<thead>
<tr>
<th>#</th>
<th>Rule Body</th>
<th>Rule Head</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>IF x is-a DemographicData AND y is-a Student AND y hasDemographics x AND x age z AND x educational_level &quot;secondary&quot; AND x gender &quot;female&quot; AND z greaterThan &quot;50&quot;^^integer</td>
<td>x computer_literacy &quot;beginner&quot;</td>
</tr>
</tbody>
</table>

\(^8\) http://www.w3.org/Submission/SWRL/
2 IF y is-a DemographicData
   AND x is-a Student
   AND z is-a TimeStudy
   AND x hasDemographics y
   AND y age w
   AND y educational_level "secondary"
   AND y gender "female"
   AND z time_for_study "much"
   AND x hasTime z

3 IF x is-a DemographicData
   AND y is-a Student
   AND z is-a ReasonsForEducation
   AND y hasDemographics x
   AND x age w
   AND x educational_level "secondary"
   AND x gender "female"
   AND z reasons_for_education "general knowledge"
   AND w greaterThan "50"^^integer

4 IF x is-a DemographicData
   AND y is-a Student
   AND z is-a ReasonsForEducation
   AND y hasDemographics x
   AND x age w
   AND x educational_level "secondary"
   AND x gender "male"
   AND z reasons_for_education "career development"
   AND w greaterThan "30"^^integer

5. Ontology-based Web Application

In order to verify the student model as described above, a web application was developed. This application enables the communication between the user through a friendly web interface and the student model ontology. According to this application, the user can import through any browser basic demographic information directly to the student model ontology.
5.1 Description of the Application

Each user can create his/her own profile as there is a simple MYSQL database with the credentials (username/password) for each individual. After having logged in, the user can import his/her basic demographic data. Below in figure 7 it is shown a screenshot of the main screen of the application.

![Student Profile - Demographic Data](image)

**Figure 7.** The main screen of the web application

Each time a new user is registered two instances of the classes *Student* and *DemographicData* respectively, are automatically created. These two instances then are connected with the object property `hasDemographicData`. Then the data properties of the
class DemographicData which represent learner’s basic demographic information, as mentioned above, take the respective values which the user has entered via the web interface (figure 7). According to the entered data, users are automatically categorized into one or more of the existing stereotypic profiles, with the implementation of the SWRL rules. In figure 8 we can see an example of how data are imported directly into the Student Model ontology. In the bottom right part, the data properties and their corresponding values appear as entered by the user. Furthermore, the user after having his demographic profile created, he/she can modify it anytime by using his username and password.

In order to deal with scalability problems, one profile (ontology) is created each time for each student. In this way we have smaller ontologies, instead of a very large document which would cost time to handle.

![Figure 8. Import of user data in the Student Model ontology](image-url)
The application is written in JSP\(^9\). For the management of the ontology and the communication with the JSP pages, the Jena API\(^10\) has been used. We have chosen Jena because it supports OWL-DL reasoning and thus we could verify that the data is consistent with the OWL restrictions in the ontology. Finally, the ontology model is stored in Jena’s SDB\(^11\). SDB is a component of Jena for RDF\(^12\) storage and query specifically to support SPARQL.

### 5.2 Evaluation of the Application

In order to test the proposed ontology and the developed web application we have gathered data from students from different schools and courses (undergraduate and postgraduate) of the Hellenic Open University.

The aim was to demonstrate that the proposed ontology is able to classify students in stereotypes in an automatic way, as these stereotypes resulted from our research.

Following, we present a case study of how the application works. Suppose that a female student named Maria Papadopoulou\(^13\), 52 years old, married and with basic education from the *Greek Civilization* course. In the basic form (figure 9) she enters her basic demographic data and then she clicks “Submit”. After the submission, Maria’s profile is created (an ontology) that corresponds to the data she has entered during the registration phase. In the following figures 10 and 11 is shown how the inference mechanisms which were incorporated in the ontology model provide the desirable result that is the classification of the student into stereotypes.

---


\(^10\) [http://incubator.apache.org/jena/](http://incubator.apache.org/jena/)


\(^12\) [http://www.w3.org/RDF/](http://www.w3.org/RDF/)

\(^13\) It’s a fake name.
In figure 10 we can see that the datatype properties (stereotypes) have taken the values according to the rules that have been incorporated into the student model ontology and we have previously seen in section 4.
For example the “computer_literacy” property has the value `beginner`, the data property “reasons-for_education” has the value `general knowledge` and the property “time_for_study” has taken the value `much` (figure 11).

At this point, is to be reminded that this model is mainly used for the initialization of the user profile.
6. Conclusions and Future Work

In this paper we proposed an ontology-based approach to model student profiles especially for distance learning students. The student profile ontology we developed can be used as an integral ITS module, while it can be easily accessed from a web-based application. The proposed approach collects the characteristics of an adult student which are considered important for an ITS in order to be fully adapted to the needs of the learner. This model is a combination of international standards in user modeling and the results of an empirical study on a group of HOU students. One of the main advantages of the proposed model is the integration of semantic rules. These rules combined with inference mechanisms classify learners into stereotypic profiles which are already incorporated in the ontology and thus produce additional knowledge. The most challenging part of our research has been the selection of the characteristics to be included in the ontology.

As a future work is the integration of more stereotypic profiles and the corresponding rules in the Student Model ontology which will allow the automatic classification of the learners into these stereotypes. Some of the stereotypes/characteristics under consideration for integration in the model are: (a) personality characteristics (e.g. extravert, thinking, etc.), (b) emotional states (e.g. happiness, anxiety, etc.) and (c) other characteristics (e.g. talkative, shy, etc.), as they are included in the General User Model Ontology (GUMO) [17]. This will increase the system’s adaptivity and provide more personalized experience to the final user within an intelligent environment.

Also, we further plan to improve the proposed application by adding more fields in order to cover the course information of a student. In this way we can extract valuable information about student's preferences and behavior during his/her academic experience. An example hypothesis that has emerged from our research is the following: “if the student has prior knowledge in the cognitive domain and has computer literacy then he/she seeks further material for his/her studies”. 
Finally, we plan to integrate the proposed model in an intelligent adaptive learning system, based on the communication between intelligent agents. This system will take as input: (a) the student profile, (b) the educational material structured in learning objects, (c) learning goals and (d) the hardware specifications. The main purpose of this system will be to provide the students with the appropriate learning object and build the learning path according to their performance and preferences.

Acknowledgment: This research described in this paper has been co-financed by the European Union (European Social Fund - ESF) and Greek National Funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) (Funding Program: "HOU").

References:


