Chapter 7

APPLYING SEMANTIC TECHNIQUES TO INTEGRATE ELECTRONIC COURSE CATALOGUES

Juan M. Santos, Martín Llamas and Luis Anido
E.T.S.E.T., Campus Universitario S/N, E-36310, Vigo (Pontevedra), Spain

Abstract: Today's technology enhanced learning landscape is characterized by a high and growing number of heterogeneous educational service providers in the international arena. This fact, intrinsically positive, raises the need of appropriate searching mechanisms that allows particular users and organizations to locate the most suitable courses for their requirements. The existence of specialized e-learning brokers or intermediaries which gather and integrate the existing educational offers (Electronic Course Catalogues) can alleviate this situation. This chapter presents the basis of an innovative brokerage system in the e-learning domain.

Key words: E-learning Brokerage; Semantic Techniques; Ontologies; Logic Rules.

1. INTRODUCTION

Currently there exists several high-quality popular search engines (like Google or Yahoo!) that provide users with results based on Information Retrieval theories; however they do not offer the appropriate support for particular contexts like e-learning. Thus, the institutions that deliver on-line courses usually promote their educational services mainly by means of the publication of electronic catalogues that are accessible through the institution's own web pages. A potential student makes use of these catalogues in order to obtain detailed information on the different courses offered and, in the case of finding some suitable, to carry out the corresponding enrolment request. To locate alternative courses to the one found is usually a complex task since the user has to manually repeat the search operation in all the institutions that he/she knows.

B. Fernández-Manjón et al. (eds.), Computers and Education: E-learning, From Theory to Practice, 77–87.

The educational services intermediation systems (or Brokers) are entities that makes it easier, on the one hand, to the different academic institutions, the publication and dissemination of electronic catalogues of offered courses and other educational resources in a common repository, and, on the other hand, to the users and potential students, the searching, comparison and location of educational resources suitable to their needs and preferences.

Brokers can collect the descriptions of the courses through standardized mechanisms as those defined in IMS-DRI (Riley and McKell, 2003) or CORDRA (Rehak et al., 2005), as well as the referring contextual information in which they are given (data about the institution, delivering tools, dates for enrolment/delivery, etc.), from affiliated academic institutions to provide high level services to people or institutions looking for appropriate online courses. The institutions, or Educational Services Providers (ESPs), can register in a Broker providing, among other data, the profile of the institution and the mechanism to access the repositories with the catalogues of the products they offer.

We are working on the development of an architecture for an innovative brokerage system in the e-learning domain which, bringing together the last standards and recommendations defined in the Learning Technologies Standardization Process and making use of the new techniques related to the emerging Semantic Web, improves the searching and location processes. The proposed architecture extends the previous works of the authors in the field (Anido et al., 2002, 2003) with semantic and inference practices.

The Business Model and the Functional Architecture of our approach is briefly discussed in Santos (2004). This chapter is mainly focused on the supporting ontology that is required, introducing several sub-ontologies about, for instance, courses and learning objects, on-line service providers, content providers, learners, etc. This ontology, named ELEARNING-ONT, provides the semantics required to let computers automatically deal with personalized intermediation in the e-learning domain.

The organization of the chapter is as follows: Section 2 describes ELEARNING-ONT, an integrative ontology for the e-learning brokerage field, Section 3 outlines the conceptual framework of the semantic brokerage architecture, Section 4 deals with the identification of logic rules required for preprocessing collected data in order to obtain significant information and, finally, Section 5 concludes and summarizes the chapter.

2. SUPPORTING ONTOLOGY

An ontology (Chandrasekaran el al., 1999) defines the terms used to describe and represent an area of knowledge (like medicine, tool manufacturing, automobile repair, financial management, etc.), including computer-usable

definitions of basic concepts in the domain and the relation amongst them. For the construction of a semantic Broker, we need the definition of a specific ontology that includes all the relevant terms required to describe all the involved entities (courses, providers, clients, e-learning platforms, etc.) and their particularities. ELEARNING-ONT is a set of interconnected OWL (McGuinness and van Harmelen, 2004) ontologies that facilitate the automatic management of the data collected and the development of intermediation services in the e-learning domain.

2.1 Development methodology

In order to identify the most suitable terms to be included in a domain OWL ontology for educational brokerage, we defined a systematic methodology. This methodology is based on the guidelines proposed by Noy and McGuinness (2000), and the recommendations described in the Unified Software Development Process (Jacobson et al, 1999).

The first stage of the development process involves the literature review and documentation of the most basic functional requirements from the client's point of view. Starting from a set of core requirements, we successively redefine the most basic "Course Search in Broker" use cases in order to capture new and different query possibilities. For each stage we apply the steps proposed by Noy and McGuinness:

1. Identification of the aim and the scope of the ontology.
2. Consider to reuse existing vocabularies (in our case, we make use of the elements defined on the data models identified by the learning technologies standardization process).
3. Enumerating the most important terms in ontology.
4. Defining the classes and their hierarchy.
5. Defining the properties of the classes.
6. Defining the features of the properties.
7. Creating instances.

In this way, the development of the ontology is an iterative process, centred on the architecture and driven by use cases, where each stage refines the previous one. As the use cases mature and are refined and specified in more detail, more of the ontology terms are discovered. In turn, this can lead to new use cases. Therefore, both the ontology and the use cases mature together.

2.2 ELEARNING-ONT description

Due to the great quantity of identified terms, the ontology is organized in a range of namespaces (or sub-ontologies). There exists a basic namespace, where fundamental concepts such as "Educational Resource", "Course" or "Educational Services Provider" are defined (c.f. Fig. 7-1). A series of sub-ontologies include the properties, with their corresponding vocabularies, that can be used to describe in detail the instances of the most basic classes:

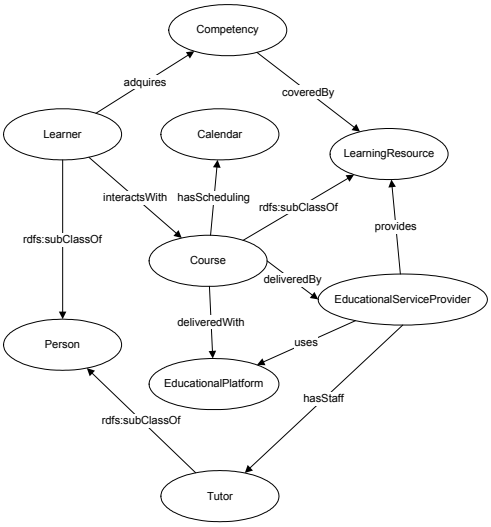

Figure 7-1. Partial view of the base ontology

- *Users Ontology:* This sub-ontology (Fig. 7-2) includes the properties and classes related directly to the characterization of the users of the brokerage system. The terms identified in this namespace have been mainly extracted from the Learning Information Package (Norton and Treviranus, 2001) and Accessibility data models (Norton and Treviranus, 2003), developed by the IMS Consortium. The first of these models identifies the necessary elements to describe the characteristics of a student, whereas the second one extends the previous model with elements that allow us to specify certain user preferences. The existence of the user's ontology allows to accomplish the searches adapted to the user needs and preferences in order to obtain more relevant results.
- *Courses and Educational Resources Ontology:* Metadata is one of the most prolific fields in the Learning Technologies Standardisation Process. Currently, the Learning Object Metadata model (Hodgins and Duval, 2002), developed jointly by several of the institutions involved in this process, is already an official standard of the IEEE. This standard,

and in particular its RDF binding, developed by Nilsson et al. (2003), has been used as the basis for the sub-ontology of ELEARNING-ONT that includes the classes and properties needed to characterize academic courses (Fig. 7-3).

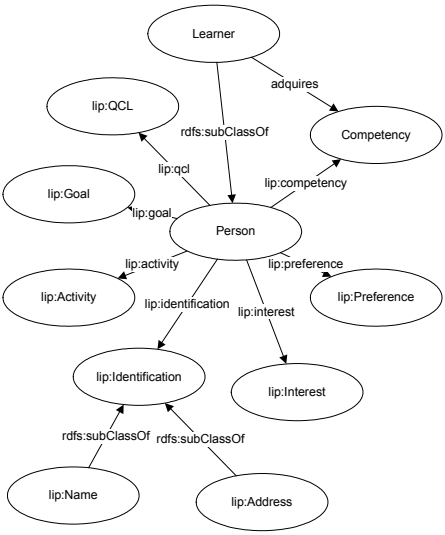

Figure 7-2. Partial view of the Users Ontology

Figure 7-3. Partial view of the Courses Ontology

- *Educational Service Providers Ontology:* Online courses are offered to students throughout e-learning platforms. An e-learning platform is a Web application that includes Internet tools and services into an enclosed space specifically configured and organized to provide learning in a convenient and satisfactory way. Many educational platform surveys have been used to elaborate the sub-ontology that allows the characterization of these applications and the terms considered to be more convenient have been taken from them. The experience of the authors related with the construction of e-learning platforms has been essential in this field. Mostly, the terms in this sub-ontology allow defining the available tools in a platform (Fig. 7-4*)*.

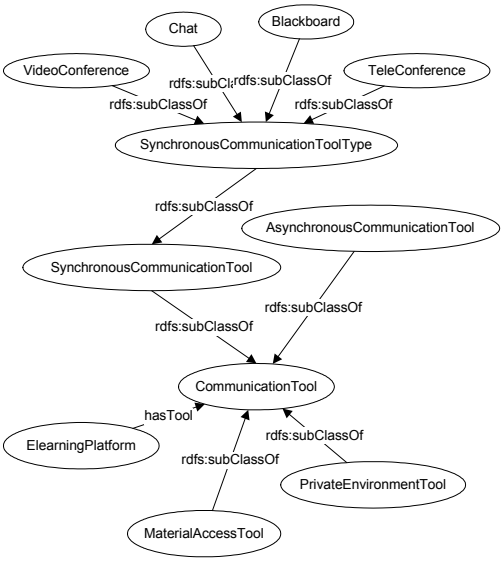

Figure 7-4. Partial view of communication tools ontology

- *Educational Platforms Ontology:* This sub-ontology gathers some terms that allow making descriptions about educational services providers. These are entities or organizations that deliver online courses throughout a particular e-learning platform. Due to the lack of standardized conceptual models in the e-learning domain related to this topic, we have taken from the e-commerce domain common schemes used that allow to describe enterprises. Particularly, our sub-ontology is based on the Enterprise Ontology (Uschold et al., 1998), developed by the Artificial Intelligence Applications Institute from the University of Edinburgh.
- *Other Ontologies and Taxonomies:* Besides the mentioned sub-ontologies, some other vocabularies and taxonomies have been used. Among them we can mention a subset of the Universal Decimal

Classification scheme, to use it as vocabulary for several of the properties defined in ELEARNIG-ONT. The DAML-Time ontology has also been imported to represent temporal concepts (for example, course calendars). Several other data models are currently under study, like ontologies that allow us to describe user's devices.

3. BROKERAGE ARCHITECTURE

Fig. 7-5 shows the functional elements of a scalable and adaptable Semantic E-learning Brokerage architecture. The proposed architecture extends the previous works of the authors in the field of semantic and inference practices. It makes use of particular ontologies (described in the previous section) and inference rules that can be refined without structural changes in the infrastructure as new statements are identified. The most important elements in the architecture are briefly described herebelow:

- *Knowledge Base:* This is the basic and core element of the brokerage system. All the information collected and inferred by the Broker is available here, both from the ESP and from the different types of clients. It is a repository where Ontologies, Inference Rules, Educational Resources and Course Descriptions, Service Provider Profiles, User Profiles and E-learning Platform Descriptions are stored.
- *Search Engine:* It is the software component that provides an API with methods for querying the Knowledge Base. Although there are many ontology query languages, currently RDQL (Seaborne, 2004) is the most used until a recommended language is issued by the W3C.
- *Inference Engine:* This component is responsible for inferring new facts from a set of previous facts taking into account additional information defined by a particular ontology and in a set of inference rules.
- *Data Collector:* It is the component that semi-automatically gathers information from the affiliated ESPs using standardized protocols (IMS-DRI, CORDRA).
- *Services:* Different services are offered by the described infrastructure. Some of them are Anonymous Searches, Personalized Searches, Notification Service, Course Annotation, Relevance Estimation Service, Taxonomy Management and Supporting Services.
- *Access Interfaces:* Different interfaces are provided to the clients in order to support different devices (PCs, PALMs, Pocket PCs, WAP devices, etc.). Likewise, Fig. 7-5 shows an access entry point for software agents. This interface consists of a set of Web Services conforming IMS-DRI.

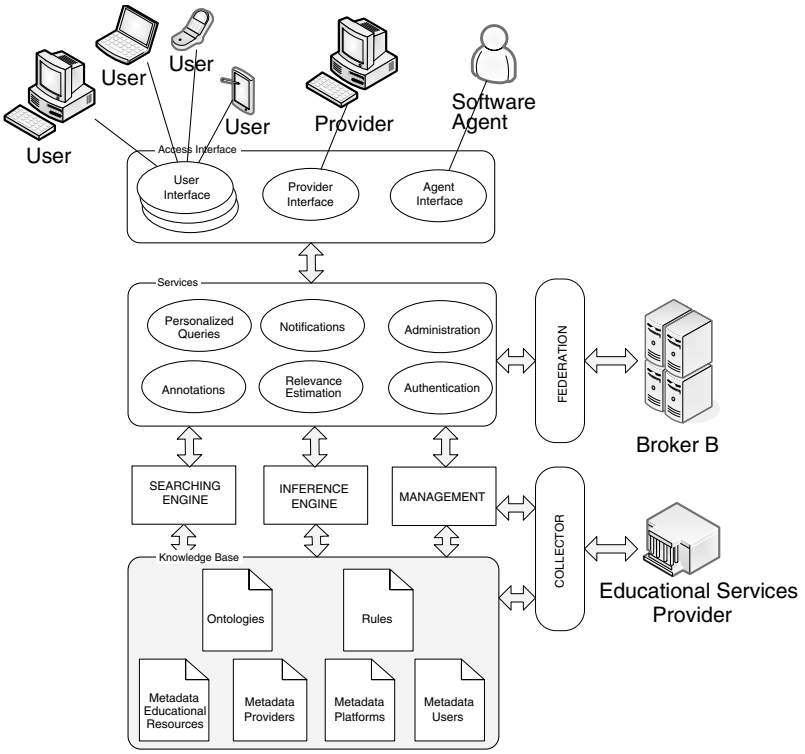

Figure 7-5. Conceptual Brokerage Architecture

4. DATA PROCESSING AND ENHANCEMENT

The data manually provided by the ESPs managers and the data automatically harvested by the Broker collector must accomplish a process of adaptation of the metadata registries obtained to the canonical format of the Broker. It must be transformed into OWL statements –or facts– that use the terms and properties defined in ELEARNING-ONT. Because the present version of IMS-DRI and CORDRA are oriented to the storage and interchange of XML-LOM descriptions, a set of transformation rules, defined on XSLT, aimed at this task, is available in the knowledge base.

The ontology-based kernel of the broker is a Knowledge Information System where facts (that describe registered courses, e-learning platforms, educational institutions and users) are stored. Logic rules can be used by an inference engine for processing and enriching the stored information and for drawing semantic conclusions. We define two basic groups of rules:

- *Semantic augmentation rules:* These rules allow making explicit knowledge that is hidden or implicit in the knowledge base. For example, the following rule (that checks the "Learning Resource Type" properties of all the elements "E" of a course "C" in order to find an element of type "simulation"):

  element(C,E), learningResourceType(E,"simulation")
  → interactivityType(C,"active")

  can be used to identify courses that are active or with a high degree of interactivity. These other two rules allow establishing a normalized "Level" property to a course and a user respectively:

  courseContext(C,"high school") → courseLevel(C,"10")
  userStudies(U,"high school") → courseLevel(C,"10")

- *Matching rules:* These rules allow inferring new knowledge by matching the preferences and characteristics of a particular user with the properties of the registered courses and their particular context. This set of rules is established in order to facilitate the searching processes. For example, the following rule (that checks the particular level "X" of a course "C" with the particular level of a user "U"):

  courseLevel(C,X), userLevel(U,X) → levelFitting(C,U)

  can be used to identify those courses that are "Level Appropriate" for a particular user.

A proper set of such explicit rules, expressed in a rule markup language, semantically complements the knowledge base of the broker. This enriched knowledge base can be queried through ontology query languages, like RDQL, to obtain relevant results for the user, providing meaningful advantages compared to traditional searching tools when we are looking for the most appropriate courses for our needs and preferences.

5. CONCLUSIONS

The application of the Semantic Web techniques in the e-learning area is being considered by multiple researchers as a valuable alternative for the improvement of several *add-on* services (e.g. in our case brokerage and location of educational objects, adaptive learning). This chapter deals with the foundation of an e-learning brokerage system that extends previous works of the authors in this field with semantic and inference practices. We have presented the basis of ELEARNING-ONT, a set of ontologies that includes the definition of the concepts, and their inter-relations, necessary to

develop brokerage services in the e-learning domain. It's an innovative proposal based on data models coming from the e-learning standardization process, allowing in this way the treatment of information which is compatible with the current trend of data and services shared between heterogeneous systems. The pre-processing of the instances of this ontology by means of suitable inference rules allows the optimization of the intermediation results in a search and location context.

ACKNOWLEDGMENTS

We want to thank MEC for their partial support to this work under grant "Metodologías, Arquitecturas y Lenguajes para la creación de servicios adaptativos para E-Learning - MetaLearn" (TIN2004-08367-C02-01). We also want to thank the DG EAC for their support through the LIFE project.

REFERENCES

Anido L., Fernández M.J., Caeiro M., Santos J.M., Rodríguez J., and Llamas M., 2002, Educational metadata and brokerage for learning resources, in: *Computers and Education*, **38**(4): 351-374.

Anido L., Rodríguez J., Caeiro M., and Santos J.M., 2003, High-level brokerage services for the e-learning domain, in: *Computer Standards & Interfaces*, **25**(4): 303-327.

Chandrasekaran B., Josephson J.R., and Benjamins V.R., 1999, What are ontologies and why do we need them, in: *IEEE Intelligent Systems*, **14**:20-26.

Hodgins W., and Duval E., 2002, *Learning Object Metadata Standard*. IEEE 1484.12.1-2002.

Jacobson I., Booch G., and Rumbaugh J., 1999, *The Unified Software Development Process*, Addison-Wesley, 1999.

McGuinness, D.L., and van Harmelen, F., 2004, *OWL Web Ontology Language Overview*. W3C Recommendation.

Nilsson M., Palmér M., and Brase J., 2003, The LOM RDF binding: principles and implementation, in: *Proc. of the 3rd Annual ARIADNE Conference*, Leuven, Belgium.

Norton M., and Treviranus J., 2001, *IMS Learner Information Package Information Model Specification*. IMS Technical Report.

Norton M., and Treviranus J., 2003, *IMS Learner Information Package Accessibility for LIP Information Model*. IMS Technical Report.

Noy N.F., and McGuinness D.L., 2000, *Ontology Development 101: A Guide to Creating Your First Ontology*. Standford-Protégé Project Report.

Riley K., and McKell M., 2003, *IMS Digital Repositories Interoperability - Core Functions information model*. IMS Technical Report.

Rehak, D.R., Dodds, P., and Lannom, L., 2005, A model and infrastructure for federated learning content repositories, in: *Proc. of the WWW'05 Workshop on Interoperability of Web-Based Educational Systems*, Chiba, Japan.

Santos J.M., Anido L., and Llamas, M., 2004, Hacia un sistema de intermediación semántico en el campo del e-learning, in: *Proc of 6º Simposio Internacional de Informática Educativa*, Cáceres, Spain.

Seaborne A., 2004, *RDQL - A Query Language for RDF*, W3C Member Submission.

Uschold M., King M., Moralee S., and Zorgios Y., 1998, The enterprise ontology, in: *The Knowledge Engineering Review*, **13**(1):31-89.