A Semantic Infrastructure to Support a Federation of Agricultural Learning Repositories

Nikos Manouselis
*Informatics Lab,
Agricultural University
of Athens, Greece
nikosm@ieee.org*

Jesús Soto Carrión
*University of Alcalá,
Spain
jesussoto@gmail.com*

Hannes Ebner, Matthias Palmér, Ambjorn Naeve
*School of Computer Science and Communication,
Royal Institute of Technology (KTH), Sweden
[hebner, matthias, amb]@csc.kth.se*

Abstract

This paper discusses how Semantic Web technologies are being adopted, specialized, and put in practice in order to set up a technical infrastructure that will facilitate sharing and reusing of learning resources. The case study is a federation of learning repositories with learning resources on agricultural topics, which is being developed in the context of the Organic.Edunet European initiative.

1. Introduction

The Semantic Web is an evolving extension of the World Wide Web (WWW) in which web content can be expressed not only in natural language, but also in a format that can be read and processed by software systems, thus permitting them to find, share and integrate information more easily [2]. Numerous applications and case studies of Semantic Web technologies (e.g. ontologies for annotating information and expressing its semantics in a machine-processable manner) have been reported during the past few years. For instance, the World Wide Web Consortium (W3C) reports on several systems that have been put in production in existing organizations, as well as a number of commercial products (http://esw.w3.org/topic/CommercialProducts). Yet, the Semantic Web technologies have not so far reached the wide public. Some of the experts in the field claim that the reason is that large-scale applications, serving the needs of large user communities, have not been delivered yet [10].

In education and training, the empowering role of Semantic Web technologies has been already acknowledged [1]. Since WWW is becoming a popular educational medium at schools, universities and professional training institutions, a prominent new stream of research on the Educational Semantic Web has been established. Research studies already report semantic-based annotation and sharing of learning resources [1,11,12]. To further illustrate this potential, there is a need for implementing state-of-the-art Semantic Web technologies in large-scale applications that involve the semantic annotation of big collections of learning resources and their access and use from existing communities of users.

This paper aims to contribute to this development by presenting such a large-scale implementation effort. More specifically, it discusses how semantic annotation and Semantic Web technologies are being adopted, specialized, and put in practice in order to set up a technical infrastructure that will facilitate sharing and reusing of learning resources. The case study is a federation of learning repositories with learning resources on agricultural topics, which is being developed in the context of a European initiative. The paper introduces this initiative and the rationale for developing a federation of learning repositories for the particular problem context. Then, it describes the architecture of the technical infrastructure and outlines the way Semantic Web technologies are involved.

2. Background

Consumer demand for food quality and safety, as well as, society's demand for more sustainable development, provide new opportunities for the agricultural sector. Public awareness of the irreversible damage done to the environment by practices that lead to soil and water pollution, depletion of natural resources, and destruction of delicate ecosystems, has led to calls for a more responsible attitude towards our natural heritage. Against this background, Organic Agriculture (OA) and AgroEcology (AE) have come to the fore as alternative agricultural philosophies towards not only producing safer agricultural products, but doing so in an environmentally sound manner.

In this light, directives such as the European Action Plan for Organic Food and Farming [3] have already identified the need for actions supporting the training and education of all stakeholders, promoting and further understanding OA and AE, and cultivating a consumer culture that will facilitate the development of an environmentally friendly products market. Large international organizations such as the United Nations' Food and Agriculture Organization (FAO, http://www.fao.org) and the International Federation of Organic Agriculture Movements (IFOAM, http://www.ifoam.org), along with non-profit associations such as the Soil Association in the UK (http://www.soilassociation.org), drive their own awareness and education initiatives for the promotion of OA methods and practices around the world. Furthermore, agricultural universities around the world have included OA-and AE-related courses in their educational programs, aiming to prepare agricultural professionals that may support and guide farmers through the selection and proper adoption of environmentally friendly methods (e.g. the AGROASIS Nordic School of AgroEcology & Ecological Agriculture, http://www.agroasis.org).

These developments have led to an increasing production of related educational content in an electronic format. On the other hand, these content resources are usually dispersed, listed in separate Web sites or individual collections, and sometimes with no clear plan for their educational exploitation. The systematic collection and categorisation of educational resources related to OA and AE, the development of an integrated online environment that will increase their use and reuse, as well as the study of educational scenarios for using this content in the context of educational programs, is a challenge that remains to be explored.

In this direction, the Organic.Edunet initiative (http://www.organic-edunet.eu) has been launched. It is a European project that is funded by the *eContentplus* Programme and involves 15 partners from 10 countries, which aims to facilitate access, usage and exploitation of digital educational content related to OA and AE. Organic.Edunet will deploy a multilingual online federation of learning repositories, populated with quality content from various content producers. In addition, it will deploy a multilingual online environment (the Organic.Edunet Web portal) that will facilitate end-users' search, retrieval, access and use of the content in the learning repositories. The project will study educational scenarios that introduce the use of the Organic.Edunet portal and content to support teaching of topics related to OA and AE in two cases of formal educational systems, i.e., high-schools and agricultural universities. Furthermore, it will evaluate project results in the context of pilot demonstrators in selected educational institutions, as well as through open validation events where external interested stakeholders will be invited.

To achieve its aims, Organic.Edunet adopts state-of-art technologies that have been developed and tested in previous research initiatives, but have yet to be proven in a real-life problem context. A characteristic example involves the implementation of Semantic Web technologies that have been previously developed in the context of the "LUISA: Learning Content Management System Using Innovative Semantic Web Services Architecture" EU project (http://www.luisa-project.eu/).

In the following section, we describe the overall architecture of the Organic.Edunet infrastructure, and how technologies such as the ones adopted from LUISA are engaged.

3. Organic.Edunet Infrastructure

The overall architecture of Organic.Edunet is illustrated in Figure 1. The main elements of this architecture are the following:

- *Learning Repository Management Module*: includes the suite of tools that the Organic.Edunet content providers will use to create a digital collection of learning resources, to describe resources with appropriate metadata, and to publish resources in their own learning repository. Overall, six learning repositories are expected to be set up by the Organic.Edunet content providers (namely the Bio@gro, ENOAT, ECOLOGICA/COMPASS, Intute, School, and Public Resources ones).
- *Learning Resource Exchange Module*: concerns the connection of the Organic.Edunet federation with other federations of learning repositories, using open standards and specifications for the exchange of search queries and the harvesting of metadata. Organic.Edunet is expected to be connected with two external federations; the Learning Resource Exchange (LRE) of the European Schoolnet (http://lre.eun.org) and the ARIADNE Foundation (http://www.ariadne-eu.org/).
- *Semantic Services Module*: it is the core of the Semantic Web technologies' application in the architecture, and supports the semantically-enabled services that the Organic.Edunet Web portal will offer, by reasoning upon a number of integrated ontologies.
- *Web Portal Module*: refers to the end-user visible parts of the whole infrastructure, allowing users (including school teachers and pupils, university teachers and students, researchers etc.) to search, locate, retrieve and access learning resources on

OA and AE throughout the whole Organic.Edunet federation.

Each module is further detailed in the paragraphs that follow.
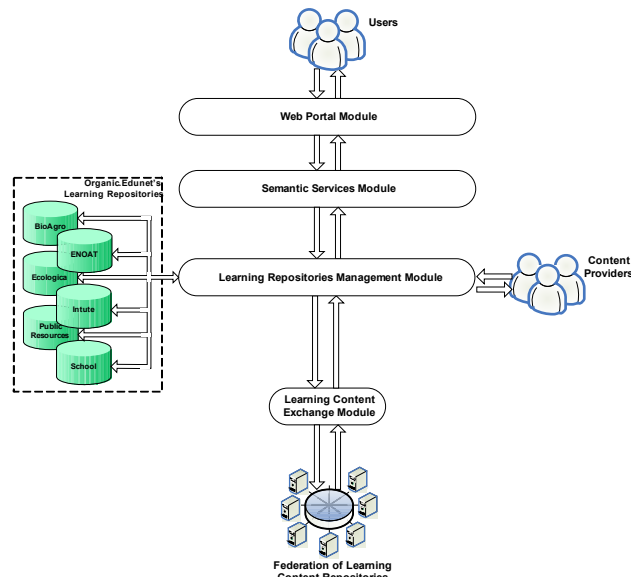
Figure 1: Overall architecture of Organic.Edunet

3.1. Learning Repository Management Module

This module deals with the way content producers organize, annotate and publish learning resources and metadata in an Organic.Edunet repository. As illustrated in Figure 2, the Organic.Edunet content producers (i.e. the organizations listed as USAMVB-FA, BCE, MOGERT, Intute, AUA and Miksike) are expected to collect and annotate their learning resources according to a multilingual application profile of the IEEE Learning Object Metadata (LOM) standard [5]. Two existing software tools are being adapted and integrated for this purpose:

- A configurable metadata editor built upon the code-library SHAME (available as Open Source at http://shame.sourceforge.net). With this code-library application programmers can develop flexible and easily extensible annotation tools for Semantic Web-based metadata. SHAME implements the Annotation Profile Model [7,8]. This model is a configuration mechanism for annotation of metadata and leaves the question of metadata standard compliance up to a metadata expert and not to the application developer.

- The electronic portfolio system Confolio (http://www.confolio.org) that allows the flexible management of folder-based repository interfaces.

The content providers will use the integration of the SHAME editor and the Confolio tool in order to upload (if desired) their resources and/or the associated metadata into a learning repository that is called SCAM (Standardized Contextualized Access to Metadata), an Open Source Semantic Web repository solution for Learning Objects [9].

The repository backend is resource-oriented and will store its metadata according to a Resource Description Framework (RDF, www.w3.org/RDF/) representation of the Organic.Edunet IEEE LOM application profile. The repository provides a range of connection interfaces, allowing the most appropriate to be chosen for each situation. An interface which exposes the repository closest to the internal representation is the REST (Representational State Transfer) [4], a resource-based software architecture building fully on top of well established standards such as the HTTP protocol. This makes it very easy to build interactive web applications on top of this interface.

The Confolio repository front-end builds on top of the REST-based web services exposed by the repository and an AJAX (Asynchronous JavaScript and XML) toolkit, which enables cross-browser compatibility and operating system independent application. The basic operations of Confolio can be separated in two groups: administrative (e.g. creation of new portfolios of learning resources) and end-user (e.g. creation of folders and description of resources using the SHAME metadata editor).

Using the Learning Repository Management Module, Organic.Edunet content providers may collect resources, annotate them using metadata conforming to the developed application profile, reviewing and approving resources, and then releasing resources for publication. Then, the metadata of the resources stored in a particular Organic.Edunet repository are (a) made available for harvesting from the Semantic Services Module and (b) made available for harvesting and/or search federation to external federations (the Bio@gro and School examples are presented in Figure 2).
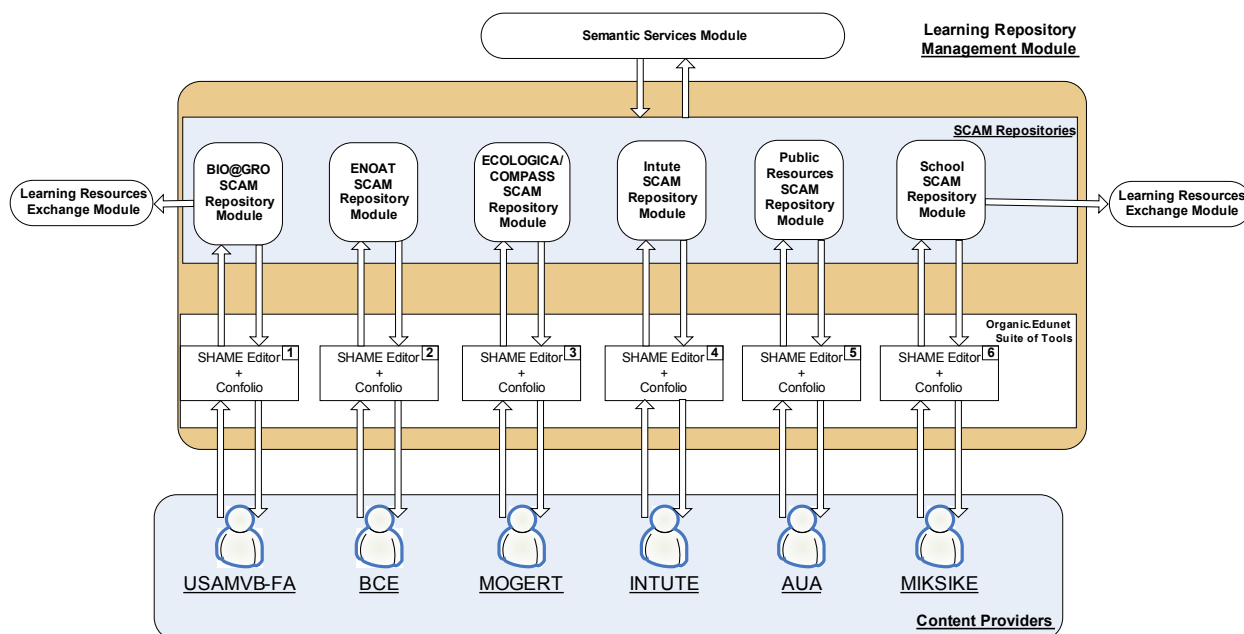
Figure 2: Overview of the Learning Repository Management Module components

3.2. Learning Resources Exchange Module

The Learning Resources Exchange Module allows for the communication of the Organic.Edunet repositories with external federations. Organic.Edunet will aim at the connection with the LRE and ARIADNE federations by adopting two widely used protocols and specifications:

- For communicating with ARIADNE: the Open Access Initiative Protocol for Metadata Harvesting (OAI-PHM, http://www.openarchives.org/OAI/ openarchivesprotocol.html) for making metadata available for harvesting from the ARIADNE services.
- For communicating with LRE: the Simple Query Interface (SQI, http://www.prolearn-project.org/lori) for serving/exchanging queries with the LRE services.

Metadata is transformed from its RDF representation into an XML representation, in order to be available for the external federations. Additional possibilities also exist for further interconnecting the Organic.Edunet repositories, due to their SCAM basis, e.g. the SPARQL Protocol and RDF Query Language - a W3C standardized query language and protocol for accessing RDF data (http://www.w3.org/TR/rdf-sparql-query/).

3.3. Semantic Services Module

The Semantic Services Module is the core engine behind the Organic.Edunet Web portal that allows offering users with semantic search capabilities. To support this, it is based on a semantic representation of the learning resources' metadata, as well as a number of ontologies that are engaged during search queries to provide reasoning capabilities. More specifically, metadata is transformed into an ontological representation inside a sub-module called LOMR.

The LOMR (standing for Learning Object Metadata Repository) is not itself a metadata repository but rather a framework which provides Web Service interfaces to any given, "real" learning object repository. LOMR instances allow developers to select the best repository implementation for a given application need, enabling specialized components, such as custom query resolvers and result composers, to benefit from the availability of different, heterogeneous LOMR instances. LOMR main features include the storage of learning object metadata in semantic format, the provision of a service-oriented interface and the import of metadata in non-semantic formats, among others.

In addition, LOMR offers Semantic Web services to the Web Portal Module, following the Web Service Modelling Ontology (WSMO, http://www.wsmo.org). WSMO provides ontological specifications for four core elements of Semantic Web Services: ontologies that provide the terminology used by other elements, goals that define the problems that should be solved by Web Services, Web Services descriptions that define various aspects of a Web Service, and finally

mediators which bypass interpretability problems. The WSML language (http://www.wsmo.org/wsml/) is used to provide a formal syntax and semantics for WSMO, since it is richer in reasoning capabilities than the OWL Web ontology language (http://www.w3.org/TR/owl-features/) recommended by the W3C. Interoperability can be easily achieved through translating WSML to OWL through open source tools that are publicly available, such as the Web Service Modeling Toolkit (WSMT, http://wsmt.sourceforge.net) and WSMO Studio (WSMO4J).

As a starting point, three ontologies are expected to be used by the Semantic Services Module. The first ontology will represent the domain area (OA and AE). It will serve all subject classification purposes, as well as allow for reasoning related to the semantics of the OA and AE concepts themselves. For example, searching for resources that have been classified using some concepts or terms related to the ones that the user has initially indicated. The popular AGROVOC (http://www.fao.org/aims/ag_intro.htm) ontology of FAO will be used as a basis. The second ontology will be a geographical one. It will help reasoning related to the geographical origin and/or coverage of resources and their associated languages. For instance, it may allow users from a particular geographical region to search for resources in languages that have been indicated as related to the particular region, even if this has not been indicated in the initial search query. The third ontology will be representing IEEE LOM. It is expected to allow reasoning related to semantics of the LOM structure itself, such as searching for information in other elements than the ones that a user has initially indicated.

In LOMR, metadata will be harvested from the individual Organic.Edunet repositories using an appropriate harvesting mechanism. As Figure 3 shows, the RDF representations stored in the SCAM repositories will be converted to the WSML representation that LOMR requires. Once all the metadata information is stored in the LOMR repository in the formal, ontology-based format, LOMR will be able to expose various functionalities through semantic Web Services (described according to WSMO), allowing a wide variety of interactions with the Organic.Edunet Web portal.
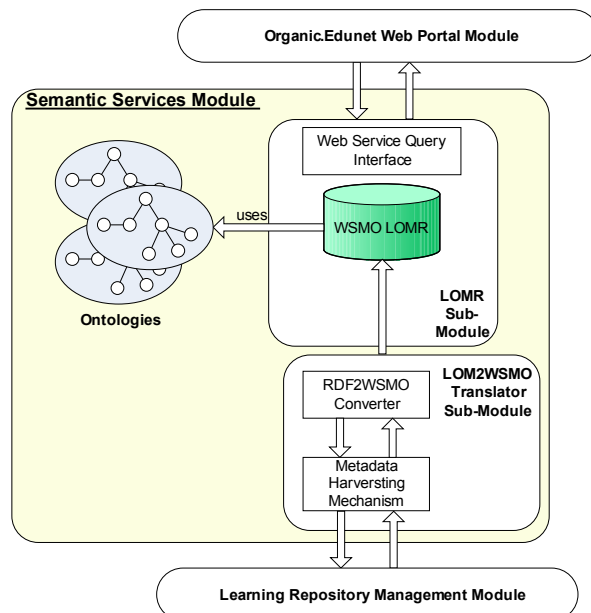
Figure 3: Overview of the Semantic Services Module

3.5. Web Portal Module

The final module of the Organic.Edunet architecture is the Web Portal one. It actually comprises the online environment that will interact with the various user roles (school teachers & pupils, university teachers and students). For this purpose, it entails a role-filtering mechanism that will allow each user category to be presented with a user interface tailored to its specific needs. Apart from allowing users to semantically search and retrieve learning resources using the Semantic Services Module, the Web Portal Module will also provide the users with the option of evaluating/rating learning resources. Multi-dimensional numerical evaluations will be stored in appropriately defined evaluative metadata [13]. Then, they will give input to a collaborative filtering mechanism that will recommend users to look at resources that other users with similar preferences liked in the past [6].

4. Conclusions

Semantic Web technologies will demonstrate their potential for education and training through real-life applications, such as semantic annotation of big collections of learning resources and semantic search and retrieval services for actual communities of users. In this direction, this paper presented a large-scale implementation effort that engages Semantic Web in order to set up a technical infrastructure that will facilitate sharing and reusing of learning resources for the agricultural domain.

5. Acknowledgment

6. References

[1] Aroyo L., Dimitrova V., "Interactive Ontology-Based User Knowledge Acquisition: A Case Study", PROLEARN Research Report 2004/4005.

[2] Berners-Lee T., " Semantic Web Road map", W3C Consortium, September 1998.

[3] EC, "European Action Plan for Organic Food and Farming", COM (2004) 415, Brussels 10 June 2004.

[4] Fielding, R. T., "Chapter 5: Representational State Transfer (REST)", Architectural Styles and the Design of Network-based Software Architectures, Dissertation, 2000.

[5] LTSC. IEEE Standard for Learning Object Metadata, 1484.12.1-2002. IEEE Learning Technology Standards Committee, 2002.

[6] Manouselis N., Costopoulou C., "Experimental Analysis of Design Choices in Multi-Attribute Utility Collaborative Filtering", International Journal of Pattern Recognition and Artificial Intelligence (IJPRAI), 21(2), 311-331, 2007.

[7] Palmér M, Enoksson F, Naeve A, LUISA deliverable 3.2: Annotation Profile Specification, 2007.

[8] Palmér M, Enoksson F, Nilsson M, Naeve A, Annotation Profiles: Configuring forms to edit RDF. In: Proceedings of the Dublin Core Metadata Conference. DCMI Conference Papers, United States, 2007.

[9] Palmér, M., Naeve, A., Paulsson, F., (2004), The SCAM-framework – helping applications to store and access metadata on the semantic web, Proceedings of the First European Semantic Web Symposium (ESWS 2004), Heraklion, Greece, May, 2004, Springer, ISBN 3-540-21999-4.

[10] Shadbolt N., Berners-Lee T., Hall W., "The Semantic Web Revisited", IEEE Intelligent Systems, 21(3), 96-101, May/June 2006.

[11] Sicilia M. A, García E. Sánchez-Alonso S, and Soto J. "A semantic lifecycle approach to learning object repositories". In Proceedings of ELETE 2005 - eLearning on Telecommunications, 2005.

[12] Soto, J., García, E. and Sánchez-Alonso, S. "Semantic learning object repositories". International Journal of Continuing Engineering Education and Life-Long Learning, Vol. 17, issue 6, pp. 432-446, 2007.

[13] Vuorikari R., Manouselis N., Duval E., "Using Metadata for Storing, Sharing, and Reusing Evaluations in Social Recommendation: the Case of Learning Resources", in Go D.H. & Foo S. (Eds.) *Social Information Retrieval Systems: Emerging Technologies and Applications for Searching the Web Effectively*, Hershey, PA: Idea Group Publishing, September 2007.