

A New Approach to the Use of Semantic Technologies in E-Learning Platforms

<http://dx.doi.org/10.3991/ijac.v9i2.5979>

Paolo Bouquet¹, Andrea Molinari^{1,2}

¹ University of Trento, Italy

²Lappeenranta University of Technology, Finland

Abstract—Semantic technologies have been studied and used in different areas of computer science. E-learning has been one of this, but the most frequent use of semantic technologies in this discipline has been in the extraction and indexing of contents, like forums, blogs, learning objects etc. The research presented in this paper aims at taking advantage of semantic technologies in a different way respect to the mainstream use that has been done in the past. This new approach refers to the use of semantic technologies in the management of the persistence layer of Learning Management Systems (LMS), i.e., where all the contents are stored. Our research follows the idea of using semantics technologies as a support, if not an entire replacement, of the backend and persistence mechanisms of LMSs. As a testbed, we will present the design and early results of the application of this approach to the persistence layer of a Virtual Communities System, where these technologies will enriched the platform to address two fundamental issues: a) entities disambiguation and identification inside the persisted objects b) adding new features to the platform without refactoring it.

Index Terms—Learning Management Systems, Semantic Technologies, Learning Objects

I. INTRODUCTION

After many different phases, e-learning seems to have reached a successful position inside every organization. In the past, the application of semantic technologies to educational settings attracted a lot of attention, in particular regarding those approaches and software tools able to enrich, categorize and retrieve learning objects.

Many researches and developments in the field of eLearning analyzed the various possibilities for existing and future e-learning frameworks to take advantage of semantic services, interoperability, ontologies and semantic annotation. Nevertheless, much of the current research seems to limit the discussions to the recurring theme of how the semantic web will enable knowledge engineers, instructors or instructional designers to construct elegant ontology-based annotations for existing web-based resources and mainly Learning Objects (LO) and to further expand metadata schemes. Minor attention has been devoted to other applications of semantics technologies to Learning Management Systems (LMS), specifically as a replacement of the backend and persistence mechanisms.

In this paper we will present a line of investigation regarding the application of semantic technologies to “Online Communities” (OC), a virtual communities platform that we are using in several projects with public and

private institutions to support educational processes. We present the design and early results of the application of semantic technologies to the persistence layer of the platform, where these technologies enrich the platform to address two fundamental issues: a) entities disambiguation and identification inside the persisted objects; b) adding new features to the platform without refactoring it, by using graph-based representation techniques to add knowledge to the existing datasets. In order to address the first item, our implementation takes advantage of the entity-centric tools developed in the Okkam EU-funded project. These tools provide a solution to uniquely and permanently identify entities (people, locations, organizations etc.) inside contents, specifically using the Entity Name System (ENS). The ENS supplies a persistent identifier, called OKKAMid, to any entity included in the knowledge base, together with advanced entity matching methods for detecting the occurrence of the same entity in different contexts and data sources. Once the entity in the LMS has been profiled, it is possible to connect any content where the different occurrences of the same entity have been mentioned inside the LMS, and also to connect any other content outside the learning platform where that entity has been named, for example web pages or social network contents.

Furthermore, we extended a domain ontology (SIOC) for the conceptual representation needs of the application, and added an RDF graph mapped onto the database, in order to add new functionalities to our virtual community platform. This allows us to take advantage of the inference processes available through a reasoner, and to substitute some parts of the business logic of the application.

To improve the semantic representation of concepts and to allow the mechanisms to interface with this knowledge base, RDF data stores have been implemented. Changing the traditional database-oriented representation of data towards a richer format (i.e. RDF statements), other than a richer representation model, gives a high level of flexibility in the definition of persistence and data representation models. RDF has been selected for data representation, while tools like Hadoop, Flink and NoSQL databases have been used and tested as a replacement of “traditional” relational databases, and a bridge towards big data scenarios. These tools and techniques have been experimented in other application fields, specifically in data integration inside the fiscal evasion domain, where different data sources are reconciled through the use of the ENS for entity disambiguation.

The paper is divided as follows. The next section will be devoted to a quick overview of how semantic technologies have been used in e-learning. The third section will

briefly present the virtual communities platform where we are experimenting the entity-centric approach to e-learning platforms. The fourth section will present a summary of the entity-centric approach promoted by the OKKAM project, and discuss how this approach can be applied for our refactoring and extension objectives. The last section will briefly present the achievements obtained with the application of semantic technologies and an entity-centric approach to our virtual communities platform.

II. E-LEARNING AND SEMANTIC TECHNOLOGIES: STATE OF ART

Most of educational institutions have recognized e-learning resources as fundamental assets for their training processes, mainly for the capabilities of delivering educational contents to participants over the Internet anytime and anywhere at competitive costs[1]. In these institutions, we can find many different implementations and customizations of available approaches (from blended to full online e-learning) and tools (platforms like LMS, technologies like videoconference, standards for learning objects metadata like LOM or SCORM [2][3]).

The maturity level of e-learning is visible also in the increasing amount of educational material available under various forms, and in the availability of many Massive Open Online Courses (MOOCs) involving people around the world. Considering the complexity of learning tasks, any new technology that can help to accelerate and improve these processes, catalyzes the attention of specialists. Semantic (web) technologies made no exception, providing a landscape where web-based information and services are understood, processed, interlinked and reused not only by humans but also by machines [4].

These are some of the suggestive perspectives about the use of semantic technologies in the e-learning field:

- the outlook to integrate learning objects enriched with metadata into an adaptive learning environment;
- the possibility of overlapping and integrate an ontological representation of the content into the functionalities of the LMS;
- the idea of integrating reasoning facilities to e-learning platform in order to infer knowledge and not previously created or complex links among educational materials.

The widest application fields of semantic technologies inside e-learning contexts have been content classification, retrieval and enrichment through the use of knowledge representation instruments, like vocabularies, taxonomies and ontologies.

Another perspective sees the semantic web as the possible implementation of a reliable, large-scale environment of machine-understandable and interoperable services that intelligent agents can discover, execute, and compose automatically [5]. Other researchers used semantic technologies to build a brand new generation of learning applications from scratch, or to enrich existing software platforms that deal with educational settings [6]. In any case, despite the excellent opportunities of combining e-Learning platforms with semantic technologies, there is no magic solution to exploit this integration: spreading the “magic powder” of semantic technologies over a LMS does not guarantee measurable improvements.

Currently, different educational standards for describing contents in learning resources exist, and a number of organizations have been involved in producing metadata standards specifically for learning technology: SCORM, IEEE LOM and IMS Learning Resource Meta-data Specification can be identified as the commonest and most robust ones[7]. Standard metadata are used by IEEE-LOM, mainly for interoperability between different LMSs but, unlike RDF based metadata, the standard only allows for a hierarchical structure. Semi-semantic metadata extend the IEEE-LOM standard with some semantic component, for example extending the relational field in the standard with a semantic net to interconnect different LO [8], or adding term associated to some pedagogical or domain ontologies [9]. Semantic metadata can be defined as “...the process of attaching semantic descriptions to Web resources by linking them to a number of classes and properties defined in Ontologies” [10]. Applications using semantic metadata rely on domain ontologies to define their metadata using RDF to express the semantics of a learning resource.

There are several advantages of using RDF over the standard metadata approach [11]: a) an RDF data model is based on the assumption of selecting metadata potentially from heterogeneous ontologies, while standard metadata are taken from a LOM-based, closed-world approach confining metadata to the particular LMS implementation; b) with RDF, complex statement can be created, thus expressing logical networks, while LOM can express simple composition of statement possibly extended through taxonomic classification; c) simple forms of inference (e.g. class inheritance, consistency check, transitivity) can be applied, this way reducing the costs of developing ad hoc solutions to implement the functions which requires them.

Annotating LO is therefore a fundamental task to guarantee and facilitate access, sharing and reuse of the learning resource. Annotation is also a keyword for the semantic world; in fact, annotated contents transform a full text to be scanned by keyword into a structured, semantically-enabled content. However, there are some obstacles to use structured learning material as a perfect knowledge base for learning activities.

Firstly, most LO have not been enriched with metadata, or have been enriched with automatic, title-related or filename-related attributes that are semantically poor and sometimes even counterproductive. Second, learning objects are not the only source of knowledge inside a LMS, and LMS platforms are not built just of learning objects. Web 2.0 tools and services, like blogs, wikis, forums, FAQs, glossaries, questionnaires etc. are most of the time very useful for the conduction of the learning process [12] especially in educational paths with a high degree of interaction among participants and instructors.

As a further element, organizations can replace their tools and platforms along time, but the investments on LO should be preserved. This means that a great attention should be paid to content and data interoperability and migration. For those materials that have been created under some standard’s umbrella, the problem should not exist, but for other contents like those created with Web 2.0 tools, the availability of an RDF-based representation simplifies the mapping process between data schemas of different e-Learning platforms [13][14] thus facilitating contents migration among different LMSs. As LOs could

be very complex multimedia artifacts, these problems could be frustrating for any interchange of educational material.

Another aspect where semantic technologies could play a fundamental role in learning settings is the addition of search capabilities to a LMS. The integration of semantic technologies is mainly devoted to get meaningful results from user queries about the knowledge base managed by the LMS itself. Parts of such a Knowledge Base that could be affected by semantic categorization could be contents, course materials, students' profiles, etc. [15].

A direction where semantic technologies frequently intercept the e-learning field is the connection with the WWW. The Semantic Web extends the categorization of existing WWW resources, allowing "computers to intelligently search, combine, and process Web content based on the meaning that this content has to humans." [16]. There have been several projects and researches that combined these three factors into e-learning systems, focusing on determining the standard architecture and format for learning environments, and this helps the integration with what has been famously illustrated in Tim Berners-Lee's "Semantic Web Stack" representation [17].

These standards, however, are trying to model the interoperability of educational information that are relevant to the educational process [18]. IMS and SCORM sequencing models define the educational activities and system implementation, together with the method for representing the intended behavior of an authored learning experience, but not the contents' knowledge in educational activities. Other authors [19] used the taxonomy of learning resources and stereotypes of teaching models for educational contents and sequences, but these aspects are heavily platform-dependent and lack standardization and reusability.

A central role is played by ontologies, here intended [20] as conceptualizations of a specific domain in terms of concepts, attributes, and relationships. Ontologies enable the representation, processing, sharing and reuse of knowledge among applications. In e-learning settings, they play a crucial role in a number of ontology-centered researches where web technology standards, such as XML and RDF(S), allow to share and reuse any web-based knowledge system [21].

The impact of the merge between e-Learning and semantic technologies is, in our opinion, deep and permanent, for reasons that we shortly presented but that have been largely discussed in different research areas [22]. Nevertheless, much of the current research seems to limit the discussions on how the semantic web will enable instructors to construct elegant ontology-based annotations for existing web-based resources, and to further expand metadata schemes [23].

Our approach is less focused on contents and more oriented to a different usage of semantic technologies, for managing unique identification of entities and increasing the information extraction from the e-learning knowledge base.

III. SEMANTIC TECHNOLOGIES AND LMS: "ONLINE COMMUNITIES" PLATFORM

The above aspects of semantic e-learning clearly emerged while expanding our virtual communities platform called "Online Communities" (OC) [24], specifically

when we started to connect the contents with external sources of information (like social networks or other web resources). OC is a collaborative environment totally designed and developed by the Laboratory of Maieutic – University of Trento (Italy) which aims at supporting cooperative processes, and teaching/learning activities in particular. Currently, Online Communities is mainly used outside the university campus, serving approximately 1000.000 users from different public and private customers against approximately 15.000 students in our University.

The core of the application is composed by some abstract entities, called "Virtual Communities", viewed as an aggregation of people to which some collaboration and communication services are available in order to obtain certain objectives [25]. (Virtual) Communities can be aggregated into larger ones, with hierarchical relationships and unlimited nesting levels, thus allowing to represent hierarchies between different types of communities (such as Faculties, Didactic Paths, Master Degrees, Courses, Departments, Organization charts etc.).

The cooperative virtual space of "OnLine Communities" is actually closed. The users participate in the system directly with their real identity. In fact, a person who enters a virtual community of our system is authorized by the community administrator and from that moment onwards he/she is automatically in contact with the people inside the community. Members of a virtual community typically have similar objectives of acquiring and sharing knowledge about specific topic. This is also why anonymity is relatively important. Lack of anonymity and control of the external accesses have origin in two explicit requirements of our first customer, the Management and Business School of the University of Trento. The exclusion of anonymity is the result of a belief that the anonymity into virtual learning environment should be banned, so that the actors cannot shirk from their responsibilities. The second circumstance (access control) stems from the will of a substantial number of teachers to block the publication on the network of their courses' Learning Objects. These choices made the system impermeable to the users' social dynamics, or to the communities existing in the social networks. Here, the role of semantic technologies could be also counterproductive, if someone wants to keep the material protected from external reference.

The architecture of Online Communities is based on five pillars: Person, Community, Service, Role and Permission. The combination of roles and permissions defines the Profile for each user. Each Community avails itself of a certain number of services, i.e., applications that enable users to communicate in synchronous and asynchronous ways, publish contents, exchange files, coordinate events, etc. Services of a community are activated by a community manager on demand, and users of a community have different rights on them.

Over the years, the system evolved into a platform for professional training oriented to lifelong learning outside academia, embracing the (social and technological) context where teaching and learning take place.

Another aspect deals with the relation between LMS and the information system of educational institutions. At the moment e-Learning platforms seem to act in a restricted circle made up of only teachers, tutors and students. In Online Communities, instead, the community is a contain-

er ready for teaching processes, but not only: research teams, recreation groups, friends, secretariats, board of directors, colleagues, anything that could represent an aggregation of people around a scope could be implemented in OC.

Given that a most detailed list of functionalities is beyond the aim of this presentation, the evolution of the platform is keeping us fully busy in studying a series of articulated functionalities:

- “traditional” services: asynchronous (forum, agenda, upload & download of learning objects, newsgroup, notice-board, classroom management, management of course pamphlets and of users, etc.) and synchronous ones (chat, streaming audio/video) and “Personalized” Services, closer to the aspects of life-long learning and “training on the job” (tutorship, training on demand, research tools with problem contextualization, semantic web, FAQ etc.)
- Integration services with external information systems (for example, the Personnel information system of the organization)
- Services for the fruition of “off-line” courses, i.e., courses already held and recorded, digitalized and made available to controlled communities of users (with the possibility to synchronize the video with slides, podcast, webcast, SCORM modules, etc.).
- Services for the management of courses delivered as Massive Open Online Courses (MOOCs). This addition meant to reuse most of the services already available in the platform, but to adapt them mainly to the “M” of MOOCs, i.e., the potential massive enrollment of people. Most of the enrollment user interface controls and widget had to be changed according to the perspective of supporting thousands of users.
- Services for the creation of evaluation test, quizzes, polls etc.
- Statistics about the users behaviour (using an internal data warehouse enriched by activity logs).
- mobile Services to support mobile learners. There are some innovative services which meet the mobility needs of the subject who wants to learn “on the move”, performing learning/collaboration activities directly through his/her mobile device (mobile phone, PDA, tablet PC, iPod, etc.).

The platform is constantly added with new services, coming from research projects, users requests and our intuitions. Among these functionalities, one is particularly frequent in users’ requests and in our “future development” discussions, i.e., propagating the visibility of files along hierarchical paths inside the communities. Let’s see a simple example: imagine a situation of communities’ relationship like the one presented in figure 1.

The file management services that are built with traditional relational technologies present several issues and limitations when requested to provide the following:

- a member of the Community “Project Management” wants to share “File 2” with its sub-communities;
- a member of the Community “Group A” wants to share “File 3” with community “Master Degree in Business Administration”;

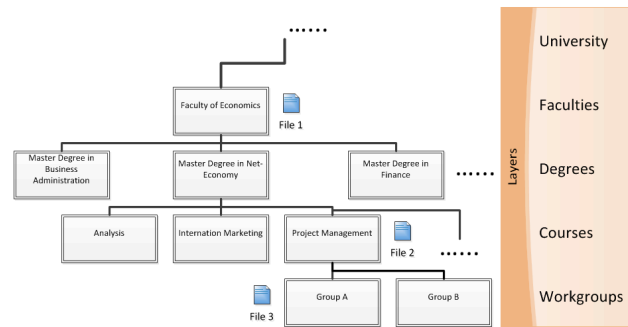


Figure 1. example of files’ inheritance in Online Communities

- the Dean of the Faculty of Economics produced a document “File 1” that should be distributed to all the communities of the Faculty.

The implementation of these functionalities have passed through different design processes, considering even the use of low-level DB-relational technologies and Object-relational mapping tools. None of these alternatives proved to be really interesting, as we should have to modify the DB and/or profoundly change the software, without considering the “brute-force” approach of replicating files in parent communities or in child communities. The three examples above are very frequent in everyday activities of OC users, and represent the need of the following mechanisms:

- *propagation*, i.e., allowing a document to be spread in sub-communities through a replication process that keep into consideration the permissions of the source and the target communities;
- *inheritance*, i.e., the previous mechanism but in the opposite direction: sub communities that want to share documents with super-communities, typically when a student produces a content that could be shared with parent communities.
- *trasversality*, in the sense of linking communities located in different branches of the communities’ hierarchy.

The problem is even more compelling, considering that in our platform, the inheritance mechanism is applicable not only to documents, but also to contents (in the sense of posts of a forum, entries in a FAQ, comments in a wiki etc.) and, most of all, to services (the user could inherit permissions on a service in sub- or super-communities, like writing permissions in a forum). In all these cases, we have two technical solutions when dealing with traditional software platforms:

- Replicating the documents where requested, with obvious drawbacks (disk usage, alignment, complexity, permissions management etc.)
- Create a (soft) link to the document needed, with a certain number of problems

All over these mechanisms, permissions and rights on services and documents must be guaranteed for every user.

Having these problems, it has been clear that the graph representation typically used in semantic technologies and representation languages could have provided a very interesting path to explore, specifically for the set of relations that are created to implement the above features.

This perfectly fit into a language of knowledge representation that allows a method of logic calculation and automatic reasoning to provide responses and take decisions. In whatever of the above conditions, traditionally engineered software application can provide solutions. In OC, most of these mechanisms are implemented and “hard coded” in the business logic and in the persistence layer of the application, but we have to face many issues:

- Query performance: we could have dramatically poor response time.
- Complexity of the overall management of these mechanisms, especially from a conceptual point of view for the user.
- User interface for rendering the inheritance relationships: for example, when a document is displayed in a list of the documents of the community, how do we represent those documents coming from other communities, or those shared with other communities?

More than this, the fascinating idea provided by semantic representation regards the idea of graph, naturally related with our idea of communities network: building an RDF graph that represents all the possible (labeled) connections among objects maintained by our platform can open new scenarios when a reasoner is applied, thus allowing the inference of logical consequences from the set of triples using appropriately rules.

From an application persistence perspective, semantic representation standards like RDF and OWL could be put aside relational representation in order to extend and empower it, allowing the implementation of new features that would be otherwise very expensive in terms of software refactoring. Apart from this economic consideration, other aspects have stimulated the development of this integration:

- the relationship among virtual communities, both hierarchical and transversal, can be more expressive than those implemented in relational databases;
- the implementation of services that use “relationship” among communities in order to inherit contents and services from related communities (files, wiki, FAQ, forums etc.);
- a graph-based navigation interface for the end-users, based on an RDF graph and interpreted through an ontology starting from the SIOC ontology [26];
- the availability of a development team with a deep knowledge of the source code of the application, that has been developed from scratch;
- the categorization of communities and contents through a tagging mechanism that allows aggregation of any object inside the platform;
- the implementation of an inferential process to access to objects/services available from parent communities, allowing the creation, for example, of “transversal wikis”, “inherited blogs”, “parent’s files”, “similar communities” and so on.

Building an RDF graph representing relevant (labeled) connections among objects can open new scenarios, especially when reasoning is applied thus allowing the inference of logical consequences in the knowledge base. Another example regards users and their management of contact lists: this is different from managing community members as a list of “friends”, or to connect people en-

rolled in the platform with the FOAF vocabulary to link people inside the platform and outside the platform through the FOAF ontology [27].

As a final example, users can accumulate many communities enrollments. These communities have contents that could be related in some way to what the user needs, but with no possibility to be inter-related with other contents in other communities. Thanks to classification and semantic representation of relationships among communities, we can create now different views and aggregation of communities (figure 2), in the future any other content that has been semantically tagged.

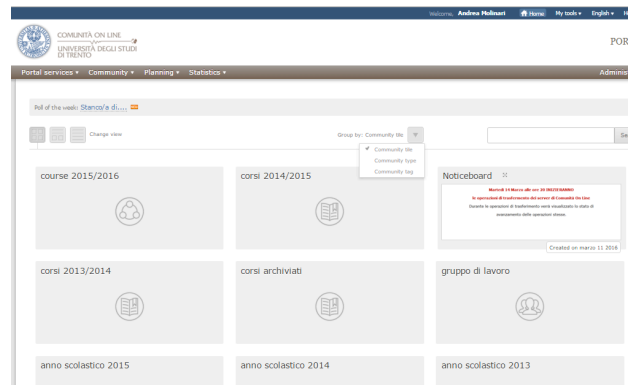


Figure 2. the tile view of tagged communities

IV. ENABLING SEMANTIC TECHNOLOGIES IN LEARNING MANAGEMENT SYSTEM

In this section, we describe how we are modifying OC with semantic technologies to achieve the results discussed above. One of the most evident advantage for the learning ecosystem was to start integrating semantic technologies with the persistence layer, in order to achieve the following results:

- linking internal contents with contents that are external to the platform;
- the creation of new “semantic-enabled” services;
- the replacement of existing services with new semantically-enriched services;
- the improvement of existing services through the use of semantic technologies
- the provision of new graph-based navigation in the platform’s entities.
- Inference on the facts represented in the semantic knowledge base

The work is inspired by the results produced during the Okkam EU project, a Large Scale Integrating Project co-funded by the European Commission between 2008 and 2010, where the core technology is the Entity Name System (ENS) [28]. The main purpose of the ENS is to provide unique and uniform names for entities for the use in information collections, so that the same name (identifier) is used for an entity, even when it is referenced in different contexts, inside and outside our platform. By reconciling entities inside the OC database, it is possible to tag “entities” inside the contents with a globally unique and persistent identifier (technically, an HTTP URI), this way providing a fundamental milestone for integrating OC with linked data available on the web.

Recognizing that information from different sources refers to the same (real world) entity is a crucial challenge in instance-level information integration, as it is a prerequisite for combining the information about one entity from different sources. The first conceptual block is the adoption of what we called an entity-centric view at the level of data and content in the persistence layer. The hypothesis is that, if all the contents inside the eLearning platform, inside the organization's information system and inside the external resources were "entity-centric", i.e., using the notion of entity for annotating and classifying data and content, the linking process among instances of the same entity through the different data sources would be straightforward. In a nutshell, this was the vision behind the Okkam vision.

The ENS is a global service (centralized or distributed, application- or private- or public- oriented) that acts as an authority for storing, recognizing and disambiguating every single entity, thanks to a unique identifier assigned to that entity. The ENS should "cut to the root the proliferation of unnecessary new identifiers for naming the entities which already have a public identifier" (inspired by the well-known Ockham's razor from the XIV century).

The first step towards the creation of a new semantic layer for OC is coupling primary keys in our database with OKKAM identifiers as provided by the ENS. We identify any relevant entity (person, place, organization, object, event, etc.) in different columns of the database tables, making sure that (i) it is recognized as an entity by the system (the "things, not strings" concept) and (ii) the same entity is always recognized as being the same entity in any type of data and context (entity resolution).

Identifying and annotating entities with OKKAM identifiers allows us to overtake several limitations of primary key usage in entity identification. While a primary key in a database represents an identifier for that record, many problems could arise from the usage of primary keys as entity identifiers:

- if the entity "XYZ" is mentioned for any reason in any content of the LMS, this will be simply taken as descriptive data, and cannot easily be connected to the other occurrences by simply annotating the text with the primary key of our table;
- identifiers generated in this way must be kept aligned with other applications of the information system where that entity is referred;
- identifiers will be invalid/useless outside the scope of the application (e.g. a LMS) where they have been created.
- identifiers should be forced by the application as a foreign key every time a relationship between our tables and other tables is established, not forgetting that this referential integrity must be reinforced by the DBMS.
- Finally, if our entity "XYZ" is present in other information systems, a) the two applications should share and preserve a common identifier, or b) the different development/ management teams need to share/exchange/align the different unique IDs, or c) the two systems will identify "XYZ" with different identifiers, thus preventing the idea of interlinked data.

This last point is particularly relevant for our argument, as it is one of our main objectives in extending our learning platform towards the semantic web. Indeed, the metadata about an entity in the ENS contains also a list of pre-existing web URIs for data about an entity on the web (mappings between an OKKAMid and other linked data URIs). So an entity in our system can be easily linked not only to any occurrence in the system itself, but also with external data which can be fetched and integrated with a simple HTTP call to other datasets about it. This is exactly paving the way towards the "web of entities" envisioned by the OKKAM project that we are embracing and presenting in this paper.

Entity identification and mapping is of course not enough to enable a full-fledged semantic application. Ontologies can be very helpful in content interpretation and integration. Also the different level of knowledge and lexicon between teacher and learner can complicate the relationship and the learning processes, also here an ontology can help a lot in sharing and transmitting understanding and knowledge. Another interesting application of ontologies in "traditional" software is to map the columns of DB tables onto concepts represented in the ontology, i.e., resolving differences among heterogeneous databases from different domains using different concepts to represent the same entity.

Using ontologies to associate unambiguous content to columns is a common approach in semantic data integration [29]. In this specific field, the ontology is a middle layer mainly used to map references to the same concepts among multiple data sources. Several techniques of schema matching have been presented, aimed at mapping elements of different database schemas that are semantically corresponding to each other in order to enable co-processing of data collected against different models.

Semantic integration of data models should also provide a way to interpret relationships between entities. The point presented in this paper sees the integration of ontologies with the Okkam entity-centric approach, useful to connect data from multiple structured and unstructured data sources referring to the same entity. What has been done so far in "okkamizing" the persistence layer of "Online Communities" platform is the extraction of primary entities from database tables, and some experiments in the analysis of forum and communities. For the purposes of the prototype we extended a very famous ontology, i.e., SIOC (Semantically-Interlinked Online Communities) to our needs. SIOC provides a Semantic Web ontology for representing rich data from the Social Web in RDF, and is commonly used together with FOAF vocabulary in order to conceptualize and present personal profiles and social networking information. We therefore used SIOC to ontologically describe some services existing in the platform, and we extended SIOC with time, events and other specific concepts available in services present in OC and not provided by SIOC or by social networks.

As a final results, we applied our idea of semantically transforming a "legacy" application into a semantic application starting from some contents of the database, mapping this part onto the ENS, adding OkkamID to entities found in this part of contents, and then creating an RDF graph with the mapped portion of the DB. This knowledge base is navigable with a browser and queryable via SPARQL.

The addition of an entity-centric approach will allow to identify entities inside contents of the platform and connect these contents with the URI referring to the same entity, thus creating a true linked data environment for e-learning. In a near future, the improvements obtained using Okkam's entity-centric approach will be quickly usable inside OC. The following are some examples of an entity-based enrichment of the knowledge base, where we can search for an Entity (not for a text) inside many different contents and services of the platform. What we did was to transfer the structure of communities and of wikis in the triple store, insert of a rule by which the reasoner infers new triples, and then prepare a web page where the prototype shows the list of Wikis and their communities before and after inference

V. CONCLUSION AND FUTURE DEVELOPMENTS

In this paper we presented a short description of experiments regarding how semantic technologies could be coupled with e-learning software, with some final considerations about the application of semantic technologies to our collaboration platform that, thanks to the metaphor of virtual communities, facilitated this kind of integration and evolution.

The early implementation revealed two of the most interesting aspects we wanted to test, i.e., the ease of implementing and/or new functionalities to a software platform when semantic technologies are wisely integrated with them. A partial objective that has been investigated regards the identification of those issues that could be generalized to any initiative aimed at extending a software platform. The idea is to put semantic technologies aside with traditional development technologies, identifying entities in contents, using persistent identifiers for storing their unique ID taking from an entities' central repository called ENS

Some further elements should be investigated, in order to have a clear vision about pros and cons of this approach, specifically regarding sustainability and investments in re-factoring e-learning applications, and in general software applications. A second item that emerged from our experiments regards performances of semantic persistence layers. Intensive tests have been conducted using the same big data-oriented technologies (Hadoop, Hbase, Flink etc.) but in different contexts, where big-data range of operation is required. As results are very encouraging, we have conducted some preliminary tests with the knowledge base available in the virtual community platform. Another items regards the exact moment where to align the database and the semantic store in case some management operations are available to the public. Finally, the potential that semantic technologies could provide to e-learning and collaboration in general is very vast, but the problem here is the usability of contents and services. Enriched by this semantic functionalities and contents, the risk is that they could become complex to be understood and therefore unusable. A great help will come from semantic tools for data visualization and navigation.

REFERENCES

- [1] Leghris, C.; Mrabet, R.; , "Cost Comparison of E-Learning Solutions," Information Technology Based Higher Education and Training, 2006. ITHET '06. 7th International Conference on , vol., no., pp.817-824, 10-13 July 2006 <http://dx.doi.org/10.1109/ithet.2006.339705>
- [2] SCORM: SCORM 2004 3rd Edition Documentation Suite, <http://www.adlnet.gov/scorm/20043ED/Documentation.aspx>, 2006
- [3] IEEE LOM: IEEE Learning Object Metadata standard <http://ltsc.ieee.org/wg12/>
- [4] Sampson, D. G., Lytras, M. D., Wagner, G., & Diaz, P. (2004). Ontologies and the Semantic Web for E-learning. *Educational Technology & Society*, 7 (4), 26-28
- [5] McIlraith, S.A., Son, T.C., & Zeng, H. Semantic Web Services. *IEEE Intelligent Systems* 16(2), 46-53, 2001 <http://dx.doi.org/10.1109/5254.920599>
- [6] Pandit, V.R.; , "E-Learning System Based on Semantic Web," Emerging Trends in Engineering and Technology (ICETET), 2010 3rd International Conference on , vol., no., pp.559-564, 19-21 Nov. 2010 <http://dx.doi.org/10.1109/icetet.2010.17>
- [7] Robson, R. Report on Learning Technology Standards. in *The World Conference on Educational Multimedia, Hypermedia and Telecommunications*. 2000.
- [8] Engelhardt, M., A. Hildebrand, D. Lange, and T.C. Schmidt. Semantic Overlays in Educational Content Networks. in *TERENA Networking Conference*. 2006. Catania, Italy.
- [9] Sancho, P., I. Martínez, and B. Fernández-Manjón, Semantic Web Technologies Applied to e-learning Personalization in <e-aula>. *Journal of Universal Computer Science*, 2005. 11(9): p. 1470-1481.
- [10] Scerri, S., C. Abela, and M. Montebello. semantExplorer: A Semantic Web Browser. In *IADIS International Conference WWW/Internet 2005*. 2005. Lisbon, Portugal.
- [11] Brooks, C. and G. McCalla, Towards flexible learning object metadata. *International Journal of Continuing Engineering Education and Life-Long Learning*, 2006. 16(No.1/2): pp. 50 - 63. <http://dx.doi.org/10.1504/IJCEELL.2006.008917>
- [12] Khamnayeve T. (2007), Applying Some Principles of Semantic Web to Alleviate Lack of Collaboration and Integration in eLearning In proceedings PICMET 2007, 5-9 August, Portland, Oregon – USA
- [13] Abels, S., Chepegin, V., & Campbell, S. (2009). Semantic Interoperability for Technology-Enhanced Learning Platforms. 2009 Ninth IEEE International Conference on Advanced Learning Technologies, 564-568. <http://dx.doi.org/10.1109/ICALT.2009.87>
- [14] Garcia, R., & Pariente, T. (2009). Interoperability of Learning Objects Copyright in the LUISA Semantic Learning Management System. *Information Systems Management*, 26(3), 252–261. Taylor & Francis. <http://dx.doi.org/10.1080/10580530903018037>
- [15] Çelik D., Elçi A., Elverici E., (2011) Finding Suitable Course Material through a Semantic Search Agent for Learning Management Systems of Distance Education, Computer Software and Applications Conference Workshops (COMPSACW), 2011 IEEE 35th Annual , vol., no., pp.386-391, 18-22 July 2011
- [16] Hitzler P., Krotzsch M., Rudolph S.. *Foundations of Semantic Web Technologies*. CRC Press, 2010.
- [17] Tim Berners-Lee. A roadmap to the semantic web. <http://www.w3.org/DesignIssues/Semantic.html>, 1998.
- [18] Adelsberger H. et al., "The Essen model: a step towards a standard learning process," <http://citeseer.ist.psu.edu/515384.html>, 2003.
- [19] Quemanda J. Simon B., "A use-case based model for learning resources in educational mediators," *Educational Technology and Society*, Vol. 6, pp. 149-163, 2003.
- [20] Gómez-Pérez, A., & Corcho, O. Ontology Languages for the Semantic Web. *IEEE Intelligent Systems* 17(1), 54-60, 2002. <http://dx.doi.org/10.1109/5254.988453>
- [21] C. Brewster et al., "Knowledge representation with ontologies: the present and future," *IEEE Intelligent Systems*, Vol. 19, pp. 72-81, 2004 <http://dx.doi.org/10.1109/MIS.2004.1265889>
- [22] Pandit, V.R.; , "E-Learning System Based on Semantic Web," Emerging Trends in Engineering and Technology (ICETET), 2010 3rd International Conference on , vol., no., pp.559-564, 19-21 Nov. 2010 <http://dx.doi.org/10.1109/icetet.2010.17>
- [23] M. A. Sicilia and E. García. On the convergence of formal ontologies and standardized e-learning. *Journal of Distance Education Technologies*, 2(4), 2004

- [24] Colazzo, L.; Molinari, A.; Villa, N.; , "Collaboration vs. Participation: The Role of Virtual Communities in a Web 2.0 World," Education Technology and Computer, 2009. ICETC '09. International Conference on , vol., no., pp.321-325, 17-20 April 2009 <http://dx.doi.org/10.1109/icetc.2009.26>
- [25] Colazzo L., Molinari A., Maresca P., Stanganelli L. (2009), Mashup Learning and Learning Communities. International Workshop on Distance Education Technologies (DET 2009), San Francisco (USA), 10-12/09/2009, p.122-128
- [26] SIOC Core Ontology Specification, 25 march 2010, <http://sioc-project.org/ontology>
- [27] FOAF Vocabulary Specification 0.98, Namespace Document 9 August 2010 - Marco Polo Edition at <http://xmlns.com/foaf/spec/>
- [28] Paolo Bouquet, Heiko Stoermer, Claudia Niederee, and Antonio Mana. Entity Name System: The Backbone of an Open and Scalable Web of Data. In Proceedings of the IEEE International Conference on Semantic Computing, ICSC 2008, number CSS-ICSC 2008-4-28-25, pages 554-561. IEEE Computer Society, August 2008.
- [29] Gómez-Pérez, A., & Corcho, O. Ontology Languages for the Semantic Web. IEEE Intelligent Systems 17(1), 54-60, 2002. <http://dx.doi.org/10.1109/5254.988453>

AUTHORS

Paolo Bouquet graduated in Philosophy in 1991 (University of Milan) and then got his PhD in Epistemology in 1997 (University of Genoa). In 2004 he became Associate Professor in Computer Science at the University of Trento, Dept. of Information Engineering and Computer Science (email: paolo.bouquet@unitn.it)

Andrea Molinari, graduated in Economics (University of Trento), is contract professor at University of Trento, Dept. of Industrial Engineering, Bolzano and Abo Akademi (Finland) (email: andrea.molinari@unitn.it)

This article is an extended and modified version of a paper presented at the the International Conference on E-learning in the Workplace 2016 (ICELW 2016), held in June 2016, at Columbia University in New York, NY, USA. Submitted 05 July 2016. Published as resubmitted by the authors 04 August 2016.