

# SEMANTIC INTEGRATION OF SOCIAL AND DOMAIN KNOWLEDGE IN A COLLABORATIVE NETWORK PLATFORM

*Luís Carlos Pacheco Soares Carneiro*

Dissertation supervised by *Prof.Dr. António Lucas Soares*  
in *INESC-Porto*

## 1. Motivation

The work described in this dissertation is contributing to the H-Know (Heritage Knowledge)<sup>1</sup>, a European research project (2009-2011) in the area of the management of old building rehabilitation, restoration and maintenance (RR&M), specially in the cultural heritage domain.

In practical terms, the project aims to develop a socio-collaborative platform where SMEs can share knowledge about restoration and maintenance activities, inducing learning and training of partners and collaboration amongst partners.

This study will focus in the **integration of semantic capabilities** for this collaborative online community platform, as a way to improve knowledge inference and retrieval of the contents produced in the platform.

## 2. Main Goals

Considering a platform supporting a collaborative network of SMEs through integrated functionalities of social networking and contents management, the specific objectives of this dissertation are:

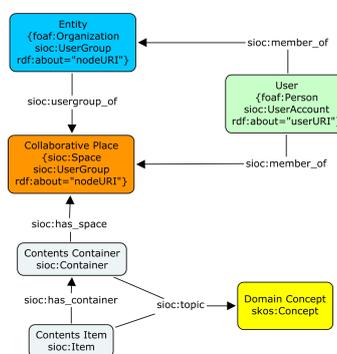
- To develop a concept and method for "semantic enabled social network" to enable the semantic integration of domain knowledge and collaboration contents;
- To develop a set of contents classification and retrieval functionalities embedded in collaboration activities (including contents created by collaboration);
- To develop a method for the creation and use of semantic information in social networking/collaboration activities.

## 3. Work Description

To build the integration of semantics in the H-Know platform and its functionalities, there are a series of decisions and analysis that were done. First of all, there was the need of analysing the Project characteristics

and background, the platform where the semantic integration is done and the domain knowledge supporting the platform. This was an important part to understand the importance of the semantic integration and to formulate "Competency Questions".

From the analysis of the problem, it was built a model (1) to semantically describe the platform.



**Fig. 1 – H-Know platform semantic conceptualization**

This model uses FOAF [1] and SIOC [2] standard vocabularies to semantically express the socio-collaborative activities, building a **platform ontology**. The domain knowledge is formalized using another standard vocabulary, SKOS [3]. Platform contents items are connected with domain concepts using the property **sioc:topic** as a bridge to **skos:Concept**.

From this model a technological architecture (Fig.2) was designed to implement it and to allow users to classify contents.

It was decided to use Sesame<sup>2</sup> as the Triple Store for the semantic metadata generated in the platform. Since Sesame API is in Java and Drupal, the framework used to build the platform, is in PHP there was the need of building a "Semantic Parser" working as an intermediate between Drupal<sup>3</sup> and Sesame for the loading and storing of metadata. To apply the model described above, the decision was to re-use existing Drupal modules which export semantic metadata, adapting them to the model and exporting the metadata to Sesame. So users can classify the contents they produce in the platform, it was built from scratch a Web

<sup>1</sup>H-Know Project Webpage <http://www.h-know.eu>

<sup>2</sup>Sesame Webpage, <http://www.openrdf.org/>

<sup>3</sup><http://drupal.org/>

<sup>4</sup><http://code.google.com/intl/en/webtoolkit/>

Application using GWT<sup>4</sup>, the "Knowledge Browser", which loads the domain knowledge ontology and gives users the chance of choosing concepts for classification when editing contents in the platform. The domain ontology is managed using the ontology editor Protégé<sup>5</sup> + SKOSed, a plugin which enables specific creation of SKOS vocabularies.

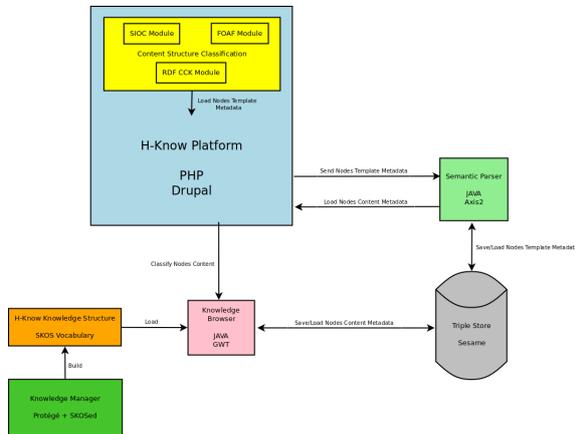


Fig. 2 – H-Know platform Semantic module architecture

To use the semantic metadata generated, there were designed some facet-browsing interfaces. A faceted classification system allows the assignment of multiple classifications to an object, enabling the classifications to be ordered in multiple ways [4]. This is an interesting way of fostering the metadata produced. We can think about facet-browsing interfaces for different purposes: search of users, entities or the most common, search of contents. In the figure 3 we can see a non-functional prototype of a contents facet-browsing interface for the H-Know platform.

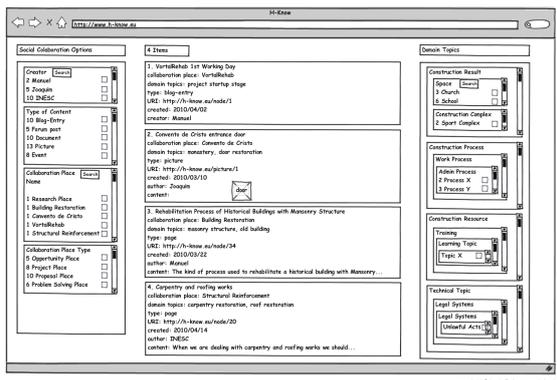


Fig. 3 – H-Know contents facet-browsing prototype

Applying filters, the user can get information resources which answer to the competency questions designed, like for example: "Which blog-entries Manuel

created about restoration processes of a church?", filtering creator for "Manuel", type for "blog-entry" and in the Domain Knowledge, Construction Result Space for "Church" and Construction Process for "Restoration".

#### 4. Conclusions

All the generated metadata, following the designed model, is stored outside of the platform in a native triple store application, Sesame. So, we have the contents created in the H-Know platform semantically described, ready to be consumed by any application that wants to take advantage of the information available in the H-Know platform.

The semantic usage of the platform generated metadata is an area that lacked of implementations during this dissertation. It was studied and conceptualized some facet-browsing interfaces to perform searches based on the semantic classifications, but those interfaces were not developed because of time limitations.

This dissertation work is obviously not perfect and complete and there are some improvements and new developments that can be made.

Thinking about the developments done in this dissertation we can think about improving the "Knowledge Browser" usability or to implement the synchronization of platform changes with the stored semantic metadata. At the moment, the semantic metadata is all exported at a time from the platform to the Triple Store.

As new developments, the facet-browsing interfaces that were conceptualized should be developed, the FOAF and SIOC metadata can be re-used in applications consuming this kind of metadata or we can think about improving the "Knowledge Browser" by presenting on it besides the information about domain concepts, platform contents according to the domain classification done by the users.

#### References

- [1] Dan Brickley and Libby Miller. FOAF vocabulary specification. <http://xmlns.com/foaf/spec/>.
- [2] Uldis Bojars and John Breslin. SIOC core ontology specification. <http://rdfs.org/sioc/spec/>.
- [3] Antoine Isaac and Ed Summers. SKOS simple knowledge organization system primer. <http://www.w3.org/TR/skos-primer/>.
- [4] M. Obitko. Translating between ontologies for agent communication. *Gerstner Laboratory for Intelligent Decision Making and Control. Series of Research Reports*, 2002.

<sup>5</sup><http://protege.stanford.edu/>