Journal of Universal Computer Science, vol. 19, no. 13 (2013), 1914-1939
submitted: 15/11/12, accepted: 28/6/13, appeared: 1/7/13 © *J.UCS*

A Semantic based Platform for Research and Development Projects Management in the ICT Domain

Carlos García-Moreno
(Indra Software Labs, Madrid, Spain
cgarciamo@indra.es)

Yolanda Hernández-González
(Indra Software Labs, Madrid, Spain
yhernandezg@indra.es)

Miguel Ángel Rodríguez-García
(Universidad de Murcia, Murcia, Spain
miguelangel.rodriguez@um.es)

José Antonio Miñarro-Giménez
(Universidad de Murcia, Murcia, Spain
jose.minyarro@um.es)

Rafael Valencia-García
(Universidad de Murcia, Murcia, Spain
valencia@um.es)

Ángela Almela
(Universidad Católica San Antonio de Murcia, Spain
aalmela@ucam.edu)

Abstract: Innovation is one of the keys to success in business and industry world, especially within the current economic context. R&D projects are a building-block in the innovation process, hence the importance of managing them efficiently. Ontologies and semantic technologies have proven highly effective at this task. Within this context, the present study explores the use of ontologies to model R&D related data and the application of semantic technologies to the building of an enhanced management system. Findings confirm the success of the system proposed, and reveal that it may bring considerable benefits to project management, such as the definition of a completely explicit information model and improved management capabilities.

Keywords: Ontology, Semantic Web, Semantic Annotation, Research and Development, Knowledge Management
Categories: M.0, M.3, M.4, M.7, M.8

1 Introduction

Innovation involves producing, assimilating and successfully exploiting a novelty in the financial and social fields in an attempt to bring new solutions to problems and to enable the fulfilment of the needs of both individuals and society [Cotec, 2004]. The need for innovation has always existed in the industry, but it has become even more pressing within the current economic context, since customers are increasingly demanding better services and products at lower prices. Thus, by being at the cutting-edge of production and marketing, customers gain a competitive advantage and the company distinguishes itself from competitors. In practice, this means improving the quality of products and reducing their price, as well as boosting productivity.

Intimately related to the concept of innovation in corporations is research and development (R&D) management, which is defined as a general set of processes and procedures for ensuring that the company succeeds in all its objectives. It usually involves several phases, namely proposal submission, project selection, approved project administration, and project deliverable dissemination. Accordingly, the sharing of project related information among different parties is crucial to the demanding decision-making tasks in R&D management. It must take place at various levels and with different knowledge backgrounds; such is the case of project applicants and reviewers, who clearly have different research interests [Liu and Ma, 2010].

Recently, certain regulations have been introduced for companies to monitor the 'quality' of their innovation activities in the industry. They issue reliable guidelines beyond the requirements laid down by other management systems, including both the effectiveness and efficiency of an R&D&I (Research, Development and Innovation) management system [UNE 166002, 2006]. Furthermore, due to its central importance, several platforms for the pooling of new ideas about this issue have been developed in the last few years [Liu and Ma, 2010]. Within this framework, semantic technologies are in an advanced stage of development [Fensel, 2002], providing a reliable basis for the challenges of organization, manipulation and visualization of data and knowledge. One of the major challenges currently posed in the area is the use of knowledge-oriented query answering to exploit the benefits of semantics.

In particular, ontologies are the fundamental technology for the semantic web. An ontology is defined as "a formal and explicit specification of a shared conceptualization" [Studer et al., 1998]. They provide a formal, structured knowledge representation, with the advantage of being reusable and shareable. In addition, they offer a common vocabulary for a domain and define the meaning of the terms with different levels of formality, as well as the relations among them. Knowledge in ontologies is mainly formalized using five kinds of components, namely classes, relations, attributes, axioms, and instances. Although their classes are usually organized into taxonomies, there are certain works in which taxonomies are considered to be full ontologies [Studer et al., 1998].

Among other applications, the use of ontologies may help overcome the limitations of the traditional methods in natural language processing. They are also relevant in the scope of some related methods such as semantic search [Lupiani-Ruiz et al., 2011], service discovery [García-Sánchez et al, 2009], information retrieval [Gharib et al., 2012] and expert finding [Lin et al., 2012]. It is also worth noting that ontologies have been applied to numerous domains, such as tourism [Jung, 2011], Robotics [Vidoni et al., 2011] or biomedicine [Ruiz-Martínez et al., 2011; Bertaud-Gounot et al., 2012]. Ontologies are thus the key to success for the Semantic Web representation. In view of that, the present study concerns the obtaining and representation of information by means of ontologies.

Specifically, we present a semantic based platform to assist in managing R&D projects. The remainder of the paper is organized as follows. A summary of the state of the art is provided in Section 2, and the components of the platform and its overall architecture are described in Section 3. Section 4 offers a use case scenario in the software development domain, and the results from the experiment are discussed in Section 5. Finally, some conclusions and future work are put forward in Section 6.

2 Related Work

Providing a platform for managing R&D projects is a paramount concern for companies and innovation managers who search for solutions to customer needs [Quelin, 2000]. Companies which are not prepared to meet the varied demands of its potential customers are at risk of losing them to better prepared competitors [Alves et al., 2005]. This need has resulted in several projects with the express purpose of facilitating knowledge management in R&D projects. For instance, [Meade et al., 2002] discuss the use of the analytic network process (ANP) method to support the selection of projects in an R&D environment. Particularly, the authors use the ANP as a model to evaluate different R&D project proposals in order to provide actors with decisions and metrics. A different approach to the management of R&D projects is the use of decision support systems (DSS), like in [Tian et al., 2005]. The authors present a method for the selection of a suitable R&D project based on DSS. This approach accounts for the organizational aspect of the decision making process.

Nowadays, the R&D management is closely related to the integration and collaboration among different stakeholders [Nobelius, 2004]. Therefore, the R&D is described as a network of actors such as competitors, suppliers, distributors, etc., focusing on collaboration within a wider system. On the other hand, Semantic Web technologies are increasingly used in projects to improve their knowledge management [Salesh, 2006]. Moreover, these technologies have been used to undertake the management of knowledge in R&D projects.

In particular, ontologies are an efficient technology for enabling semantics-driven knowledge systems [Razmerita et al., 2003]. Hence, ontologies can be applied to knowledge management systems to provide them with conceptual models. They are being incorporated in real-world project solutions; for instance, [Maedche et al., 2003] present an integrated knowledge management architecture for companies. Similarly, [Bullinger et al., 2005] describes an ontology-based project for managing innovation projects, trying to improve the competitiveness of companies by facilitating the acquisition of knowledge of new innovative products. Thus, this ontology-based system supports employees in the knowledge acquisition process. Other systems such as the one presented in [Colomo-Palacios et al., 2010] are centred in providing a system that enables the management of R&D intermediary firms based on semantic technologies and open standards.

One of the burning issues in the field is the efficient management of R&D projects in multilingual environments. Projects in this line should support the sharing of R&D information among researchers from different nationalities who use different languages. Thus, some projects such as [Liu and Ma, 2010] covers this issue by defining an ontology-based methodology to the development of R&D project management systems with multilingual supports. This methodology is validated during the development of a management system which supports three different languages to facilitate R&D information sharing among users in the Innovation and Technology Commission of the Hong Kong SAR Government.

On the other hand, semantic approaches for expert finding has gained momentum and some approaches have arisen lately. For example, in [Alemán-Meza et al., 2007], a framework of existing vocabularies (FOAF, SIOC and SKOS) for expert finding is proposed. Other approaches make use of the Linked Open Data that allows querying the whole Web like a huge database, thus surpassing the limits of closed data sets, and closed online communities [Stankovic, 2010; Stankovic et al., 2010]. Finally, other approaches are based on the analysis of multiple features of natural language documents of the company [Abramowicz et al, 2011; Zhu et al., 2010].

The framework presented in this work provides a decision support system for R&D industrial project managing. This decision support system allows identifying previous R&D projects and the best team for a new R&D project. Our platform has been designed to reuse and take advantage of existing vocabularies and ontologies such as FOAF, DOAP and ResumeRDF. On the other hand, it uses natural language features of the documents of the semantic descriptions in order to integrate all the best practices of the current approaches.

3 Platform architecture

The architecture of the proposed system is shown in Figure 1. The system is based on previous works [García-Moreno et al., 2011; Valencia-García et al., 2011] and is composed of 4 main modules: the ontology repository, the semantic annotation module, the semantic indexing module and the semantic inference engine. To put it briefly, the system works as follows. All the research projects, staff and material resources of the companies are semantically described and inserted in the ontology repository. For each resource, the semantic annotation module extracts all the semantic annotations with respect to a domain ontology. Next, the semantic indexing module creates a semantic index based on the Vector space model. Finally, when a new project is introduced, the semantic inference engine suggests previous related projects and the best team and material resources which are to be assigned to this new R&D project. Next, these components are described in detail.

3.1 Ontology repository

The main objective of this module is to store all the ontologies used by the system. Three different types of ontologies exist: R&D ontology, the domain ontologies, and the RDF (Resource Description Framework) based semantic profiles. Below, a more detailed description of these components is provided.

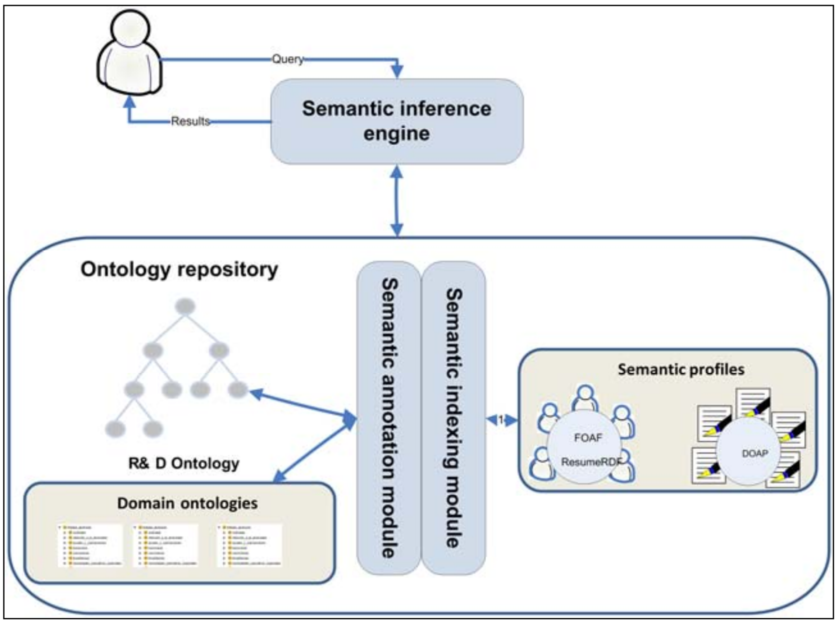

Figure 1: System architecture.

3.1.1 R&D ontology

For this work, an R&D ontology based on the Spanish standards on innovation has been developed. The standards used to define this ontology are the following:

- "UNE 166006:2006 - R&D&I Management: Technological Surveillance": it describes the information related to the activities involved in the process of monitoring the technology [UNE 166006, 2006].
- "UNE 166001:2006 - R&D&I Management: Requirements of an R&D&I project": it contains information that is relevant when implementing an R&D&I project [UNE 166001, 2006].
- "UNE-EN ISO 9001:2000 - Systems Quality Management: Requirements": it emphasizes the possibilities concerning compatibility between the different management systems of a company.
- "INDRA Software Labs Quality Manual": it was required to implement the use case.
- "UNE 166005:2004 - UNE 166002:2002 application Guide": it indicates how to apply the standard through the implementation of an innovation management system by experts [UNE 166005, 2004].

The R&D ontology has been designed from scratch and it has been implemented using the second version of the "Web Ontology Language", OWL 2 [Grau et al., 2008]. In particular, OWL 2-DL has been used, which is based on description logics (DL) and supports a number of important automatic DL inference services. These services can be provided by DL reasoners including Pellet2, HermiT or Racer [Sirin and Parsia, 2004] and fulfill the following tasks:

- To ensure ontology consistency, as far as contradictory features are concerned.
- To check concept satisfiability to know whether it is possible for a class to have any instances. If a class is unsatisfiable, then defining an instance of the class will cause the whole ontology to be inconsistent.
- To provide a classification service to compute subclass relations among all the named classes, to create the complete class hierarchy. The class hierarchy can then be used to answer queries such as getting all the subclasses of a given class or just the direct ones.
- To find the most specific classes to which individuals belong in order to compute their direct types.

3.1.2 Domain ontologies

The ‘domain ontologies’ repository stores relevant knowledge of the application domain where the platform is going to be applied. In this case, the domain is information and communications technology (ICT), therefore an ontology describing this domain has been developed.

The first problem about defining this domain is the organization of the information on the interest domain, especially as regards its structuring. Both problems were solved so that other information source could be reused to structure and organize the information domain. This information source is the Wikipedia website, which has been our main information provider. Wikipedia is one of the biggest information containers on the Internet. We have focused on the design of an extraction process to facilitate the recovering and organization of the information of the domain at issue. Figure 2 shows an extract of the knowledge base used by the platform with some of the generic terms of the ontology, where the organization of the information can be seen.

The generic concepts shown are related to the domain in question (ICT) to build and organize our knowledge base. As can be seen in Figure 2, they are organized in a tree-shaped structure where the concepts closest to the root node (owl:Thing) describe generic information on our domain, whereas the elements in the leaves are more specific concepts within the domain.

3.1.3 Semantic profiles

In this section, the semantic profiles for representing the main resources of the R&D projects are described, such as the material and human resources involved in the company. These semantic profiles are extracted from the information included in the databases of the company. For this purpose, an application for extracting the semantics from these databases has been implemented. These semantic profiles are described in detail below.

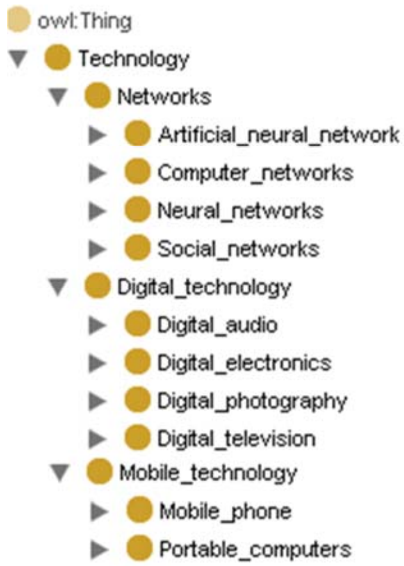

Figure 2: An extract of the domain ontology

3.1.3.1 Semantic descriptions of R&D projects

Description of a Project (DOAP[1]) is an RDF schema and XML vocabulary to describe software projects, in particular open-source projects. It was created to convey semantically-enriched information associated with open-source software projects.

The vocabulary in DOAP is insufficient for describing an R&D project. For example, DOAP profiles do not account for all the possible roles which can be played by the participants in an R&D project. Therefore, a new DOAP-based RDF schema has been created by extending the vocabulary in order to improve the semantic description of an R&D project in an organisation. In Figure 3, an example of a DOAP-based description of a project is offered. There, the doap_ext prefix is used for referencing the name gap of the DOAP extended vocabulary, including the *project_coordinator*, *project_participant*, and *project_manager* roles.

3.1.3.2 Semantic descriptions of human resources

Socio-semantic networks involve agents who produce, manipulate and exchange knowledge or information. They are groups of distributed knowledge creation and processing or knowledge communities: communities of scientists, free software developers, "wiki" contributors and webloggers, to name but a few [Rotha and Cointet, 2010]. In a company, these agents can play many different roles in an R&D project within the ICT domain, such as project manager, documenter, tester, or programmers. It is worth noting that one actor can play different roles within the

[1] http://trac.usefulinc.com/doap/

scope of an R&D project. For example, the project manager can also belong to the testing team.

The main objective of this ontology is to represent the roles and competences of the participants in an R&D project. The roles played by the personnel within the company, along with other semantic information, are represented by means of FOAF (Friend-of-a-Friend), ResumeRDF and DOAP-based semantic profiles. FOAF is an RDF representation of people and their friends' personal information. ResumeRDF, on the other hand, is an ontology developed to convey the information contained in a résumé, such as business and academic experience, skills, publications, certifications, etc. [Bojärs and Breslin, 2007]. Finally, the extended DOAP-based profile defined in the previous section allows R&D projects to be semantically described.

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
        xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
        xmlns:foaf="http://xmlns.com/foaf/0.1/"
        xmlns:doap="http://usefulinc.com/ns/doap#"
        xmlns:doap_ext="http://www.innovation-labs.com/ns/doap_ext#">
        <doap:Project>
          <doap:name>SITIO</doap:name>
          <doap:shortdesc xml:lang="en">
The SITIO project deals with the problem of the increasing number of Business Processes
available as Services on the Web, which gives rise to a new challenging problem: the
integration of heterogeneous applications. Semantic Technologies deal with adding machine-
understandable and machine-processable metadata to Web resources through its key-
enabling technology: ontologies. Ontologies are a formal, explicit and shared specification of
a conceptualization. The breakthrough of adding semantics to Business Processes for SaaS
leads to an architecture to integrate various Information Systems from the perspective of
emerging scenarios which benefit from the use of such technologies.
</doap:shortdesc>
          <doap_ext:project_coordinator>
                <foaf:Organization>
                      <foaf:name>Indra Software Labs S.L</foaf:name>
                </foaf:Organization>
          </doap_ext:project_coordinator>
          <doap_ext:project_participant>
                <foaf:Organization>
                      <foaf:name>Universidad de Murcia</foaf:name>
                </foaf:Organization>
                <foaf:Organization>
                      <foaf:name>Universidad Carlos III de Madrid</foaf:name>
                </foaf:Organization>
          </doap_ext:project_participant>
          <doap_ext:project_manager>

                <foaf:Person>
                      <foaf:name>Carlos García Moreno</foaf:name>
                </foaf:Person>
  </doap_ext:project_manager>
        </doap:Project>
</rdf:RDF>
```

Figure 3: An example of a DOAP description of an R&D project

Semantic profiles represent the experience of users in topics related to the domain and R&D ontologies. The personnel's semantic profile includes:

- The development projects in which they are –or have been– involved (DOAP).
- Their work experience in other companies (ResumeRDF).
- Other personal information (FOAF).

In Figure 4, an example of a ResumeRDF-based description of a worker is provided. There, the *cv* prefix is used for referencing the name gap of the ResumeRDF vocabulary, *foaf* prefix is used for specifying the name gap of the FOAF vocabulary, and *doap* and *doap_ext* prefixes are used for referencing the name gap of the extended DOAP vocabulary.

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
   xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
   xmlns:foaf="http://xmlns.com/foaf/0.1/"
   xmlns:doap="http://usefulinc.com/ns/doap#"
   xmlns:doap_ext="http://www.innovation-labs.com/ns/doap_ext#">
   xmlns:cv=" http://rdfs.org/resume-rdf/#">
  <cv:CV>
   <cv:aboutPerson>
    <foaf:Person>
     <foaf:name>Carlos García Moreno</foaf:name>
    </foaf:Person>
   </cv:aboutPerson>
   <cv:hasEducation>
    <cv:Education>
     <cv:studiedIn>Universidad Carlos III de Madrid</cv:studiedIn>
     <cv:degreeType>Computer Science MSC</cv:degreeType>
    </cv:Education>
   </cv:hasEducation>
   <cv:hasWorkHistory>
    <cv:WorkHistory>
     <cv:employedIn>Indra Software Labs S.L.</cv:employedIn>
     <cv:jobTitle>Innovation manager</cv:jobTitle>
     <cv:isCurrent/>
    </cv:WorkHistory>
   </cv:hasWorkHistory>
   <doap_ext:workedIn>
     <doap:Project>
      <doap:name>SITIO</doap:name>
     <doap:Project>
   </doap_ext:workedIn>
  </cv:CV>
</rdf:RDF>
```

Figure 4: An example of a semantic description of a worker.

3.1.3.3 Semantic descriptions of material resources

The main objective of this ontology is to represent the machines and materials that have been used in an R&D project. The information of these material resources in the company along with other semantic information are represented by means of the GoodRelations and DOAP-based semantic profiles. GoodRelations is an OWL lightweight ontology for annotating offerings and other aspects of e-commerce on the Web; it provides a standard vocabulary for denoting things, machines and other kind of products [Hepp, 2008]. On the other hand, the extended DOAP-based profile defined in the previous section allows R&D projects to be semantically described and enables the representation of the project in which the material products have been used.

Figure 5 shows an example of the semantic description of a server. As can be seen, this server has been used under the projects SITIO and SONAR.

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
   xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
   xmlns:foaf="http://xmlns.com/foaf/0.1/"
   xmlns:gr="http://purl.org/goodrelations/v1#"
   xmlns:doap_ext="http://www.innovation-labs.com/ns/doap_ext#">
  <gr:ProductOrService>
    <gr:name>Dell Server 54</gr:name>
    <gr:category>Server (computing)</gr:category>
    <doap_ext:usedIn>
       <doap:Project>
         <doap:name>SITIO</doap:name>
       <doap:Project>
    </doap_ext:usedIn>
    <doap_ext:usedIn>
       <doap:Project>
         <doap:name>SONAR</doap:name>
       <doap:Project>
    </doap_ext:usedIn>
   </cv: ProductOrService>
</rdf:RDF>
```

Figure 5: An example of a semantic description of a server.

3.2 Semantic annotation module

This tool receives both the domain ontologies and the semantic profiles of R&D projects. Then, the natural language texts of these profiles are extracted and, by using a set of natural language processing (NLP) tools, a semantic annotation for the natural language texts in accordance with the domain ontologies is obtained. This module is based on the methodology presented in [Valencia-García et al., 2008], and the semantic profiles are annotated with the classes and instances of the domain ontologies by going through the process described below. First, the most relevant linguistic expressions are identified by means of statistical approaches based on the syntactic structure of the text. Then, for each linguistic expression, the system tries to determine whether it is an individual of a class within the domain ontology. Apart from it, the GATE framework2 has been employed. GATE is an infrastructure for developing and deploying software components which process human language. GATE helps scientists and developers in three ways: (i) by specifying an architecture, or organizational structure, for language processing software; (ii) by providing a framework, or class library, which implements the architecture and can be used to embed language processing capabilities in diverse applications; (iii) by providing a development environment built on top of the framework made up of convenient graphical tools for developing components.

3.3 Semantic indexing module

In this module, the system retrieves all the annotated knowledge from the previous module and tries to create fully-filled annotations with this knowledge. This step is based on the work presented in [Castells et al., 2007]. The annotation of each semantic profile is stored in a database and has a weight assigned, which reflects how relevant the ontological entity is for the profile meaning. Weights are calculated by using the TF-IDF algorithm [Salton and McGill, 1983], which uses Equation 1:

$$(tf - idf)_{i,p} = \frac{n_{i,p}}{\sum_k n_{k,p}} * \log \frac{|P|}{N_i} \qquad (1)$$

where $n_{i,p}$ is the number of occurrences of the ontological entity i in the semantic profile p, $\sum_k n_{k,p}$ is the sum of the occurrences of all the ontological entities identified in the semantic profile p, $|P|$ is the set of all the semantic profiles, and N_i is the number of all profiles annotated with *i*.

In this scenario, the natural language texts in the semantic profiles described in the previous section are analysed. For each profile, an index is calculated based on the adaptation of the classical vector space model presented in [Castells et al., 2007]. Each profile is represented as a vector in which each dimension corresponds to a separate ontological concept of the domain ontology. The value of the dimension of each ontological concept is calculated as follows (see Equation 2):

$$(v_1, v_2, ..., v_n)_p \ where\ v_i = \sum_{j=1}^{n} \frac{tf - idf_{j,p}}{e^{dist(i,j)}} \qquad (2)$$

where *dist(i,j)* is the semantic distance between concept *i* and concept *j* in the domain ontology. This distance is calculated with the taxonomic (subclass_of) relations of concepts in the domain ontology. Thus, the distance between a concept

[2] http://gate.ac.uk/

and itself is 0, the distance between a concept and its taxonomic parent or child is 1, and so on and so forth.

In Figure 6, an excerpt of the ICT ontology is shown; specifically, some concepts related to cloud computing and semantic web technologies are provided.

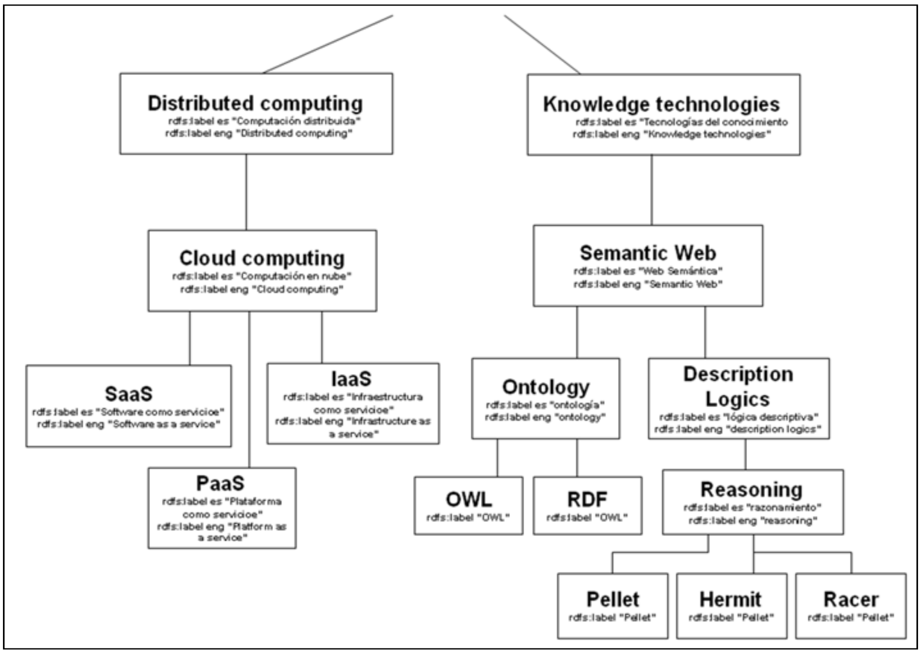

Figure 6: An excerpt of the ICT ontology.

Let us suppose that there were the text descriptions of the two R&D projects provided below (Figures 7 and 8):

The main objective of SITIO Project is the application of several emerging paradigms (**Software-as-a-Service**, **Semantic technologies**, Business Process Management, **Cloud Computing**) to build a new **Software-as-a-Service** platform that will include **Semantic technologies**, oriented to interoperability and to the reduction of costs, with the resulting profit for the software industry.

Figure 7: Text description of the first R&D project.

The goal of SONAR project is focused on the development of a corporate financial search engine based on **knowledge technologies** and using **semantic web** elements. Basically, system collects data from two sources, public (Internet) and private (Intranet) corporate information, and merges these data into an **ontology** repository and allows the access to these data in an intelligent way, through a **reasoning** engine, which uses the **semantic content** and a natural language processor that helps users to do the searches as easy as possible.

Figure 8: Text description of the second R&D project.

In Figures 7 and 8, the annotations extracted by the semantic annotation module are presented in bold. The semantic annotation module matches the concepts of the ontology with the text.

Once the annotations have been obtained, the indexes of each profile have to be calculated using the equations described in (1) and (2). In Table 1, the most important components of the vector for each project description are shown. As shown below, concepts not appearing in the text descriptions obtain a score due to the semantic distance with other concepts in the ontology.

Document	**Concepts**	**tf-idf**	**Concept dimension value**
1	SaaS	0,049147754	0,053667866
	Semantic technologies	0,000816327	0,000816327
	Cloud Computing	0,012286939	0,030367387
	Distributed computing	0	0,011171537
	IaaS	0	0,011171537
	PaaS	0	0,011171537
2	Knowledge technologies	0,007167381	0,009582609
	Semantic web	0,004761905	0,009202003
	Ontology	0,00358369	0,006483918
	Reasoning	0,00358369	0,004763408
	Description logics	0	0,004525172
	OWL	0	0,0023853
	RDF	0	0,0023853
	Pellet	0	0,00175236
	Hermit	0	0,00175236
	Racer	0	0,00175236

Table 1: Some of the concept dimensions of the vector space model

The score of the relational database concept dimension is higher than the one obtained just with the tfidf algorithm. The use of semantic distance between concepts in the ontology improves the results and reduces ambiguity. Thus, the main idea is that, semantically speaking, the closer the concepts, the higher the values of their dimensions.

3.4 Semantic inference engine

The semantic inference process benefits from the semantic content in the system, as explained below. Figure 9 shows the way in which the semantics associated to the R&D projects is represented within the system. Each person is described by means of the FOAF and ResumeRDF vocabularies, and each material resource is described by means of the GoodRelations vocabulary. The projects in which these people and material resources have been involved are described by means of the DOAP vocabulary and, finally, the annotations of each project are linked to the domain ontology.

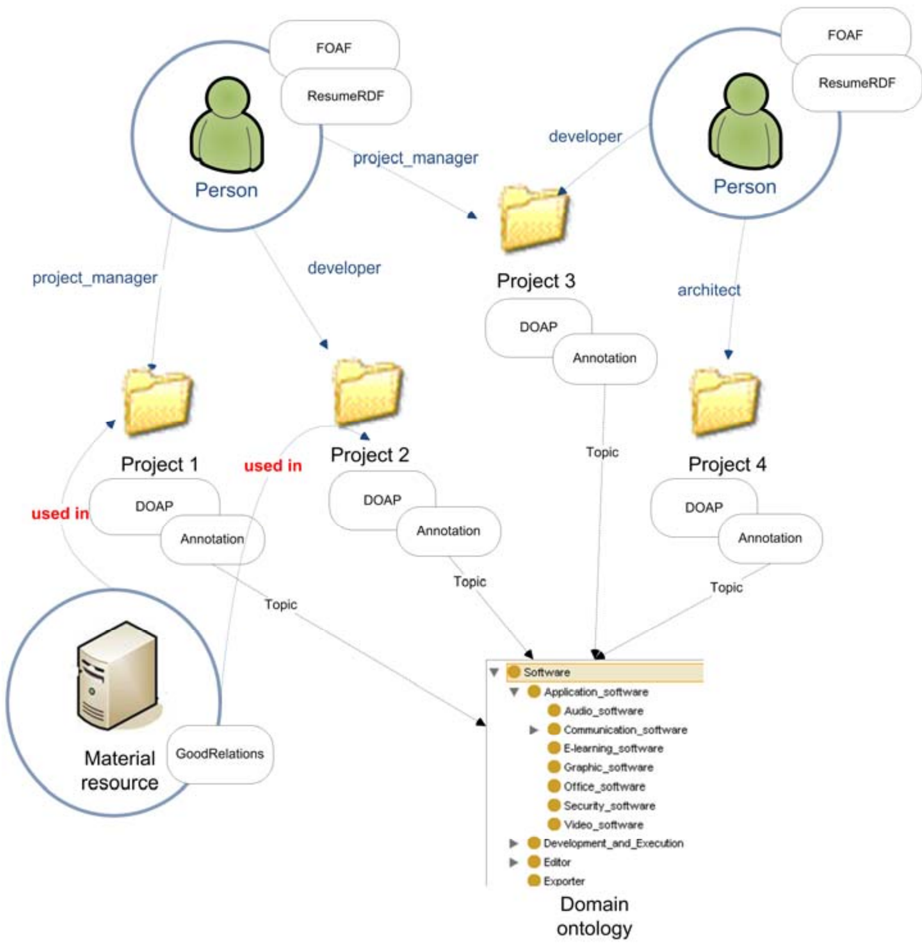

Figure 9: Connections between the vocabularies of the system.

Once a new R&D project is introduced and semantically described using DOAP-based profiles (see Figure 9), and the relevant annotations of the project have been obtained, the semantic inference engine suggests a set of similar projects already existing in the knowledge base, as well as the people and material resources that should be used in the new project. A list of the previous related projects and the people and material resources involved are rated and sorted by relevance, according to the semantic profiles and the similarity to the current project. These steps are explained in detail below.

3.5 Suggesting similar R&D projects

As stated above, an R&D project is defined by a DOAP-based profile and a semantic index obtained from annotations of the domain ontology. In order to determine the similarity between two software projects, semantic similarity functions based on previous works [Valencia-García et al., 2011] have been incorporated into the system.

Intuitively, the evaluation of the correlation between two projects must consider the sum of the similarity of their DOAP-based profile and the similarity between the semantic indexes of them. Given two projects, the similarity between them is calculated as shown in the following equation:

$$projectSim(P_i, P_j) = \alpha_1 * DOAPSim(P_i, P_j) + \alpha_2 * \cos Similarity(sem\_index(P_i), sem\_index(P_j)) \quad (3)$$

where $\sum \alpha_i = 1$ and $0 \le \alpha_i \le 1$.

Here, DOAPSim(P*i,* P*j*) is the quantitative value of the similarity of the DOAP-based profiles and the $\cos Similarity(sem\_index(P_i), sem\_index(P_j))$ is the cosine similarity to the semantic indexes of each project. The following equation is used for determining the ontological similarity to each part of the semantic profile (ontology) of a project.

$$DOAPSim(P_1, P_2) = \sum \max_{\substack{i=1..|O_1| \\ j=1..|O_2|}} (concSim(conc(P_1, i), conc(P_2, j)) \quad (4)$$

where conc(P,i) obtains the concept i of the semantic profile of a project P. Given two DOAP-based semantic profiles, P1 and P2 in the above formula calculate the conceptual similarity with each concept of the first profile in relation to the second one. Conceptual similarity between concepts is calculated using the following equation:

$$concSim(c_i, c_j) = cp_1 * conc\Pr ox(c_i, c_j) + cp_2 * propSim(c_i, c_j) \quad (5)$$

where $\sum cp_i = 1$ and $0 \le cp_i \le 1$.

Thus, conceptual similarity is calculated as the weighted average of conceptual proximity (concProx) and the similarity of the set of properties of each concept (propSim). Specifically, conceptual proximity calculates how near the concepts are in the ontology. Here, nodes stand for the number of concepts between Ci and Cj through the shortest common path, that is, through the closest common taxonomic parent concept. Providing that there is no common parent, proximity scores 0. The function is defined by the following equation.

$$conProx(c_i, c_j) = \begin{cases} 1 - \dfrac{dist(c_i, c_j)}{totalNodes} & (anc(c_i) \cup \{c_i\}) \cap (anc(c_j) \cup \{c_j\}) \\ 0 & otherwise \end{cases} \quad (6)$$

where *totatNodes* stands for the total number of concepts in the ontology, and *anc(c)* returns the taxonomic parents of concept c.

Finally, *propSim* accounts for the similarity between the sets of properties associated to the respective concept (see Equation 7). It is calculated by using the similarity measurement presented in related work, e.g. [Fernández-Breis et al., 2009].

$$propSim(c_i, c_j) = \frac{common(c_i, c_j)}{common(c_i, c_j) + nonCommon(c_i, c_j)} \quad (7)$$

The factor *common(ci, cj)* refers to the number of properties (relations and attributes) shared by both concepts, and *nonCommon(ci,cj)* is analogous, but considering the set of attributes and relations which do not appear in both concepts.

On the other hand, the cosine similarity calculates a similarity value between the vector which represents the semantic index of the current project P and each stored project P'. Cosine similarity is used for this purpose (see Equation 8):

$$sim(P, P') = \cos\theta = \frac{P \bullet P'}{|P| \bullet |P'|} \quad (8)$$

The first vector (P') is calculated through Equation 8 for each R&D project, and the second vector (P) is the one created from the concepts extracted from the current project. The θ symbol is the angle separating both vectors; it represents the degree of similarity between them.

3.5.1 Detecting people and material resources involved in similar projects.

Once the similarity between the current R&D project and the previous ones has been calculated, and assuming that the similarity factor of one or more of these previous projects reaches a given threshold, it is necessary to retrieve the human and material resources involved in those similar projects. In order to obtain the participants and their role in a given project, the following SPARQL query (see Figure 10) is entered into the ontology repository.

```
PREFIX foaf:<http://xmlns.com/foaf/0.1/>
PREFIX doap:<http://usefulinc.com/ns/doap#>
PREFIX doap_ext:<http://www.innovation-labs.com/ns/doap_ext#>
SELECT ?person ?role
WHERE { doap:projectX ?role ?person.
    ?person a foaf:Person
    }
```

Figure 10: SPARQL query for obtaining people who have worked in the project.

This query obtains the workers involved in *projectX* and the role played (developer, project_manager, etc). The score associated to each person is the sum of the project similarity score for all the projects in which the person has been involved. Equation 9 is considered:

$$peopleScore(A,P) = \sum_{hasParticipated(A,P_i)} projectSim(P,P_i) \quad (9)$$

where P_i refers to the semantic description of the previous software development projects carried out by the company.

The results are then sorted by the score of each person. For each evaluated person, the system provides information about the projects (s)he has worked on, and the parts which (s)he has played in those projects.

On the other hand, to obtain the information about the material resources of a project, the SPARQL query shown in Figure 11 is entered.

```
PREFIX foaf:<http://xmlns.com/foaf/0.1/>
PREFIX doap:<http://usefulinc.com/ns/doap#>
PREFIX doap_ext:<http://www.innovation-labs.com/ns/doap_ext#>
PREFIX gr:< http://purl.org/goodrelations/v1#>
SELECT ?material
WHERE { doap:projectX doap_ext:used ?material.
    ?material a gr:ProductOrService
    }
```

Figure 11: SPARQL query for obtaining material resources that have worked in the project.

This query obtains the material resources used in *projectX*. The score associated to each material is then calculated as the sum of the project similarity score for all the projects in which the material has been used. Equation 10 is considered:

$$materialScore(A,P) = \sum_{hasBeenUsed(A,P_i)} projectSim(P,P_i) \tag{10}$$

where P_i refers to the semantic description of the previous software development projects carried out by the company.

3.6 Semantic Search Engine

Company managers in charge of selecting the human resources and materials for R&D projects can be assisted by tools for identifying the resources which best suit a new project. The system described here uses a semantic search engine to conduct the search over the semantic profiles registered in the system.

Current approaches for semantic search use formal ontology-based queries, such as SPARQL [Prudhommeaux, 2005]. These languages return tuples with the ontology values matching the queries. In this work, the platform makes use of the SPARQL language. For instance, in Figure 12 an example of an SPARQL query for obtaining all the people that have worked in semantic web R&D projects and the roles they have played in those projects is provided.

```
PREFIX foaf:<http://xmlns.com/foaf/0.1/>
PREFIX doap:<http://usefulinc.com/ns/doap#>
PREFIX doap_ext: <http://www.innovation-labs.com/ns/doap_ext#>
PREFIX tic:<http://www.innovation-labs.com/ns/tic#>
SELECT ?person ?role ?project
WHERE { ?project ?role ?person.
      ?person a foaf:Person.
      ?project a doap:Project
      ?project doap:annotation tic:SemanticWeb
     }
```

Figure 12: An example of a SPARQL query.

4 Use case: ICT domain and Indra Software Labs

Indra is a global company of technology, innovation, and talent, leader in high value-added solutions and services for sectors such as transport and traffic, energy and industry, public administration and healthcare, finance, insurance, security and defence, telecom and media, to name but a few. Indra operates in over 100 countries and has 30,000 employees worldwide who share their knowledge and experience in different sectors and countries to find innovative solutions to the challenges faced by clients. Indra is the European company that most invests in R&D in its sector.

Specifically, Indra Software Labs is an Indra’s subsidiary specialized in software development. It is characterized by the division of work between its various development centres and the creation and implementation of various methodologies, processes and tools for high productivity and quality products.

This platform has been implemented in Indra Software Labs, and an ICT ontology has been developed under the scope of project IMAN, funded by the "Centro de desarrollo tecnológico e industrial CDTI".

The ICT ontology has been developed using the ontology evolution approach presented in [Rodríguez-García et al., 2012]. This approach allows using the Wikipedia website as the information provider to extract new concepts and taxonomic relationships to be included in the ontology. For this purpose, a manual built ontology comprising a reduced number of concepts was developed. Specifically, it contained around 250 concepts and 400 taxonomic relationships. This ontology was developed by means of the tool presented in [Rodríguez-García et al., 2012]. Table 2 illustrates some metrics concerning the ICT ontology.

	ICT ontology
Classes	1747
Datatype properties	39
Object properties	155
Subclass_of relationships	5379
Classes with Multiple Inheritance	259
Max. Depth of Class Tree	11
Min. Depth of Class Tree	2
Avg. Depth of Class Tree	5,9
Max. Branching Factor of Class Tree	113
Min. Branching Factor of Class Tree	1
Avg. Branching Factor of Class Tree	6

Table 2: Details about the ICT ontology.

Once the ontology was built, managers of the company were required to input data on some previous software projects of the company and to describe them from a semantic perspective; this information was stored in the ontology repository. Virtuoso repository [OpenLink Software, 2009] has been used to implement the ontology repository. Around 40 R&D software projects were semantically described and introduced into the system. An average of 5 people per project participated, so a total of 100 different semantic descriptions of people were introduced into the system. The FOAF- and ResumeRDF-based profiles were inputted into the ontology repository through a web application. Once the participants of the projects had been introduced, the semantic description of each project was manually defined. Then, the semantic annotations and indexes were calculated automatically from the natural language descriptions in the project profile using the modules explained in section 3. For this purpose, a web application for managing these DOAP-based profiles in the ontology repository was developed.

Once the semantic information about the previous projects and the personnel who had worked in those projects had been introduced, the experiment took place. This experimental evaluation aims at the assessment of the usefulness of the semantic inference engine of the proposed platform. Accordingly, 10 new R&D projects within the context of ICT were proposed. For each project, an expert generated the corresponding semantic description, which was subsequently included into the repository. Then, a board of human resources managers manually selected the most related old projects and the best possible team from the company's staff to participate in every new R&D project. At the same time, the team advisor module was asked to perform the same task automatically. The results obtained by the inference engine were then compared to those achieved by the human experts.

The evaluation of the results has been done through the precision, recall, F1-measure and MAP scores. The precision score (see Equation 11) is obtained by dividing the amount of proposals put forward both by the system and in the manual process by the total amount of proposals automatically made. The recall score (see Equation 12) is obtained by dividing the amount of proposals put forward both by the system and in the manual process by the total amount of items (projects and experts) manually selected. The F_1 score (see Equation 13) is the weighted average of precision and recall; the F_1 score reaches its best value at 1 and the worst score at 0. Finally, the Mean Average Precision (MAP) defined in Equation 14 has been used.

$$precision = {correct\ suggestions} \Big/ {total\ suggestions} \quad (11)$$

$$recall = {correct\ suggestions} \Big/ {number\ of\ relevant\ items} \quad (12)$$

$$F_1 = 2 * {precision * recall} \Big/ {precision + recall} \quad (13)$$

$$MAP = \frac{\sum_{p=1}^{P} \Pr ecision(p)}{P} \quad (14)$$

where Q is the number of different projects and the precision function is calculated using Equation 11.

The R&D projects used in this experiment can be classified into four main categories: software engineering projects, human computer interaction projects, ambient assisted living projects and projects on the Internet of services. The results obtained are shown in Tables 3, 4 and 5.

5 Discussion of the experiment

Tables 3 and 4 provides information on the performance of the obtained results for both project and team suggestions in terms of precision, recall and F-measure. The results shown seem promising.

The difference in the precision values of the similar projects suggestion is significant. For example, projects 9 and 10 have obtained a low precision score (i.e., 71% and 65%, respectively). This is due to the fact that the company is short of experience in this type of projects and there are no similar previous projects in the system. On the other hand, projects 7 and 8 have obtained a very good precision score (i.e., 88% and 82%, respectively). This is because the company has been centred in the development of ambient assisted living technologies and there is a broad background of previous projects in the ontology repository.

		Projects suggested		
Project type	**Project**	**Precision**	**Recall**	**F1**
Software engineering (4)	**1**	0,75	0,66	0,70
	2	0,84	0,70	0,76
	3	0,81	0,72	0,76
	4	0,78	0,65	0,71
Human computer interaction (2)	**5**	0,69	0,62	0,65
	6	0,72	0,70	0,71
Ambient assisted living (2)	**7**	0,88	0,76	0,82
	8	0,82	0,64	0,72
Internet of services (2)	**9**	0,71	0,61	0,66
	10	0,65	0,59	0,62
Total		0,77	0,67	0,71

Table 3: Precision, recall and F_1 measure of the experiment, project suggested.

		Team suggested		
Project type	**Project**	**Precision**	**Recall**	**F1**
Software engineering (4)	**1**	0,71	0,63	0,67
	2	0,76	0,64	0,69
	3	0,74	0,68	0,71
	4	0,73	0,58	0,65
Human computer interaction (2)	**5**	0,68	0,60	0,64
	6	0,73	0,66	0,69
Ambient assisted living (2)	**7**	0,77	0,68	0,72
	8	0,70	0,53	0,60
Internet of services (2)	**9**	0,65	0,58	0,61
	10	0,63	0,55	0,59
Total		0,71	0,61	0,66

Table 4: Precision, recall and F_1 measure of the experiment, team suggested.

Regarding the recall score of the similar projects suggestion, projects 9 and 10 have obtained the worst results, namely 61% and 59%, respectively, because a very low number of similar projects were obtained with the projects on the Internet of services. The best results have been obtained in the ambient assisted living projects, due to the fact that this kind of projects are very specific and less ambiguous, and almost all of them were suggested. Finally, the method achieved the best results in the F-measure in project 7, obtaining a score of 82%.

The overall results for project suggestions seem promising, obtaining a precision score of 77%, a recall score of 67% and an F-measure score of 71%.

The results obtained for the team suggestion are worse than the ones obtained for the project suggestion. For example, projects 9 and 10 obtained the worst precision because the company does not have many employees that have a broad background on technologies of the Internet of services. The best precision scores were obtained in ambient assisted living and software engineering projects, since the company employees have wider experience in these topics. Specifically, the best precision score is obtained in project 7, namely 77%. A regards the recall score, the team suggested for project 8 obtained the worst result, because some workers with considerable experience in ambient assisted living technologies were not suggested. The bad results obtained in the projects on the Internet of services are due to the fact that the company does not have many people with experience in this type of projects. The overall results for team suggestions are quite good, obtaining a precision score of 71%, a recall score of 61% and an F-measure score of 66%.

On the other hand, Table 5 shows the MAP measure of the experiment. As can be seen, the best results for both projects and team suggestions are obtained in ambient assisted living projects (i.e. 85% and 74%, respectively).

It is difficult to compare different approaches described in the literature, because none of the software applications are available and the experimental sets are unavailable. Indeed, the examples used in similar works differ significantly in content and size; concepts and relations are treated differently; domains, relations and attributes are different; errors are not treated uniformly, etc. Besides, some of the most similar works do not provide any statistical metric of the results [Alemán-Meza et al., 2007; Colomo-Palacios et al., 2010; Stankovic et al., 2010; Abramowicz et al., 2011]. A fair comparison of similar methods would require the usage of the same testing examples and the same ontology.

In fact, the only comparison can be done with expert finding systems such as the one presented in [Stankovic, 2010], where the F-measure scores obtained were between 49% and 71%. As can be seen, our approach improves this results, obtaining an F-measure between 59% and 72%.

Some works use the MAP measure for evaluating its performance. For example, the work presented in [Zhu et al., 2010] obtains the best MAP score of 65,64%, while our approach obtains a MAP score of 71%.

Project type	**Projects suggested MAP**	**Team suggested MAP**
Software engineering	0,80	0,74
Human computer interaction	0,71	0,71
Ambient assisted living	0,85	0,74
Internet of services	0,68	0,64
Total	**0,77**	**0,71**

Table 5: MAP of the experiment.

6 Conclusions and future work

Innovation is one of the keys to success in nowadays business world, especially within the current economic context. In most cases, innovation can lead to an increase in profits by improving the quality of the outcome and decreasing production costs. R&D projects are a building-block in the innovation process; they form a sound basis for companies to meet their business and strategic objectives. Traditional R&D management systems are fraught with problems derived from the need to share heterogeneous data among different departments and levels.

Ontologies and semantic technologies have proven highly effective at capturing, defining, sharing and reusing the knowledge on a specific domain. In this work, we propose the use of ontologies to model R&D related data and the application of semantic technologies to build an enhanced R&D management system. The main benefits derived from adding semantics to our R&D management system are:

1. The definition of a completely explicit information model, since the existence of ontologies serves as a reference for communication (both among people and computers) and helps improve data quality and consistency.
2. Improved search capabilities, obtained by describing R&D projects and participant profiles in terms of a well-defined and formal domain model.
3. An improved management of information, to be achieved by means of the exploitation of search and inference capabilities.

As further work, we aim to extend the scope of the application to cover the whole R&D&I management lifecycle. The ultimate goal is that R&D&I management in companies is conducted in an integrated, holistic way. All the departments in the company should be involved in this process, which must be cyclical as far as use and generation of knowledge are concerned. Therefore, the entire lifecycle must be considered, from the initial brainstorming to the analysis of the results of the finished projects.

Acknowledgements

This work has been supported by the Centro de Desarrollo Tecnológico Industrial (CDTI) through project IMAN (IDI-20110605) and by the Spanish Ministry of Economy and Competitiveness and the European Commission (FEDER / ERDF) through project SeCloud (TIN2010-18650).

References

[Abramowicz et al., 2011] Abramowicz, W., Bukowska, E., Kaczmarek, M., Starzecka, M.: "Semantic-enabled Efficient and Scalable Retrieval of Experts", Proceedings of Third International Conference on Information, Process, and Knowledge Management (eKNOW) (2011)

[Alemán-Meza et al., 2007] Alemán-Meza, B., Bojars, U., Boley, H., Breslin, J.G, Mochol, M., Nixon, L. JB., Polleres, A. and Zhdanova, A.V.: "Combining RDF Vocabularies for Expert Finding"; In E. Franconi, M. Kifer, and W. May (Eds.): ESWC 2007, LNCS 4519 (2007), 235–250

[Alves et al., 2005] Alves, J., Marques, M. J., Saur, I. and Marques, P.: "Building creative ideas for successful new product development"; 9th European Conf. on Creativity and Innovation, Lodz, Poland, (2005).

[Bertaud-Gounot et al., 2012] Bertaud-Gounot V., Duvauferrier R., Burgun A.: "Ontology and medical diagnosis"; Informatics for Health and Social Care, 37, 2 (2012), 51–61.

[Bojärs and Breslin, 2007] Bojārs, U., Breslin J.G.: "ResumeRDF: Expressing skill information on the Semantic Web"; 1st International ExpertFinder Workshop, Berlin, Germany, January (2007).

[Bullinger et al., 2005] Bullinger, H., Warschat, J., Schumacher, O., Slama, A., Ohlhausen, P.: "Ontology-Based Project Management for Acceleration of Innovation Projects"; Lecture Notes in Computer Science, 3379, (2005).

[Castells et al., 2007] Castells, P., Fernández M., Vallet D.: "An adaptation of the vector-space model for ontology-based information retrieval"; IEEE Transactions on Knowledge and Data Engineering, February, 19, 2 (2007), 261–272.

[Colomo-Palacios et al., 2010] Colomo-Palacios, R., García-Crespo, A., Soto-Acosta, P., Ruano-Mayoral, M., Jiménez-López, D.: "A case analysis of semantic technologies for R&D intermediation information management"; International Journal of Information Management, 30, 5 (2010), 465-469.

[Cotec, 2004] Cotec Foundation for Research and TechnologyWhite paper: "The Innovation Spanish System"; (2004).

[Fernández-Breis et al., 2009] Fernández-Breis, J.T., Castellanos-Nieves, D. Valencia-García, R.: "Measuring Individual Learning Performance in Group Work from a Knowledge Integration perspective"; Information Sciences, 179 (2009), 339-354.

[Fensel, 2002] Fensel, D.: "Ontologies: A Silver Bullet for Knowledge Management and Electronic Commerce"; Springer, (2002)

[García-Moreno et al., 2011] García-Moreno, C., Hernández-González, Y., Hernández-Alcaraz, M.L., García-Sánchez, F., Valencia-García, R.: "Adding semantics to Research and Development management"; In DCAI 2011, Salamanca, Spain (2011).

[García-Sánchez et al., 2009] García-Sánchez, F., Valencia-García, R., Martínez-Béjar, R., Fernández-Breis, J. T.: "An ontology, intelligent agent-based framework for the provision of semantic web services"; Expert Systems with Applications, 36, 2 (2009), 3167-3187.

[Gharib et al., 2012] Gharib, T.F., Badr, N., Haridy, S., Abraham, A.: "Enriching Ontology Concepts Based on Texts from WWW and Corpus"; Journal of Universal Computer Science, 18, 16 (2012), 2234-2251.

[Grau et al., 2008] Grau, B., Horrocks, I., Motik, B., Parsia, B, Patel-Schneider, P., Sattler, U.: “OWL2 the next step for OWL”; Web Semantics: Science, Services and Agents on the World Wide Web, 6, 4 (2008), 309-322.

[Hepp, 2008] Hepp, M.: “GoodRelations: An Ontology for Describing Products and Services Offers on the Web”; Proceedings of the 16th International Conference on Knowledge Engineering and Knowledge Management (EKAW2008), Acitrezza, Italy, September 29 - October 3, Springer LNCS, 5268, (2008), 332-347.

[Jung, 2011] Jung, J.J.: “Exploiting multi-agent platform for indirect alignment between multilingual ontologies: A case study on tourism business Expert Systems with Applications”; 38, 5 (2011), 5774-5780.

[Lin et al., 2012] Lin, F., Sandkuhl, K., and Xu, S.: “Context-based Ontology Matching: Concept and Application Cases”; Journal of Universal Computer Science, 18, 9 (2012), 1093-1111

[Liu and Ma, 2010] Liu, O., Ma, J.: “A multilingual ontology framework for R&D project management systems”; Expert Systems with Applications, 37 (2011), 4626-463.

[Lupiani-Ruiz, 2011] Lupiani-Ruiz, E., García-Manotas, I., Valencia-García, R., García-Sánchez, F., Castellanos-Nieves, D., Fernández-Breis, J.T. Camón-Herrero, J.B.: “Financial news semantic search engine”; Expert systems with applications, 38, 12 (2011), 15565-15572.

[Maedche et al., 2003] Maedche, A., Motik, B., Stojanovic, L., Studer R., Volz, R.: “Ontologies for Enterprise Knowledge Management”; Intelligent Systems, IEEE Computer Society, 18, 2 (2003), 26-33.

[Meade et al., 2002] Meade, L.M., Presley, A.: “R&D project selection using the analytic network process”; IEEE Transactions on Engineering Management, 49, 1 (2002), 59.

[Nobelius, 2004] Nobelius, D.: “Towards the sixth generation of R&D management”. International Journal of Project Management, 22, (2004) 369-375.

[OpenLink Software, 2009] OpenLink Software: “Virtuoso: Universal Server Platform for the Real-Time Enterprise”; (2009).

[Quelin, 2000] Quelin, B.: “Core competencies, R&D management and partnerships”; European Management Journal, 18, 5 (2000), 476.

[Prudhommeaux, 2005] Prudhommeaux, E. Seaborne, A.: “SPARQL Query Language for RDF”; February (2005), http://www.w3.org/TR/2005/WD-rdf-sparql-query-20050217/

[Razmerita et al., 2003] Razmerita, L., Angehrn, A. and Maedche, A.: “Ontology-Based User Modeling for Knowledge Management Systems”; Lecture Notes in Computer Science, Springer. 2702, (2003), 148-148.

[Rotha and Cointet, 2010] Rotha, C., Cointet, J. P.: “Social and semantic coevolution in knowledge networks”; Social Networks, 32, (2010), 16–29.

[Rodríguez-García et al., 2012] Rodríguez-García, M.A., Valencia-García, R. and García–Sánchez, F.: “An Ontology Evolution-Based Framework for Semantic Information Retrieval”; Lecture Notes in Computer Science, Springer 7567 (2012), 163-172.

[Ruiz-Martínez et al., 2011] Ruiz-Martínez J.M, Valencia-García R., Fernández-Breis J.T., García-Sánchez F., Martínez-Béjar R.: “Ontology learning from biomedical natural language documents using UMLS”; Expert Systems with Applications, 38, 10 (2011), 12365-12378.

[Salesh, 2006] Salesh, M.: "Semantic Web Knowledge Management"; Advances in Systems, Computing Sciences and Software Engineering, Springer Netherlands, (2006), 101-106.

[Salton and McGill, 1983] Salton, G., McGill, M. J.: "Introduction to modern information retrieval"; McGraw-Hill. ISBN 0070544840, (1983).

[Sirin and Parsia, 2004] Sirin, E., Parsia, B.: "Pellet: An OWL DL reasoned"; Description Logic Workshop (DL 2004) (2004).

[Stankovic, 2010] Stankovic, M.: "Open Innovation and Semantic Web: Problem Solver Search on Linked Data"; In Proceedings of International Semantic Web Conference (ISWC) 7th-11th Novebmer, Shanghai, China (2010)

[Stankovic et al., 2010] Stankovic, M., Wagner, C., Jovanovic, J., & Laublet, P.: "Looking for Experts ? What can Linked Data do for You?"; In Proceedings of Linked Data on the Web 2010, Raleigh, NC. (2010)

[Studer et al., 1998] Studer, R., Benjamins, V. R., Fensel D.: "Knowledge engineering: Principles and methods"; Data Knowl. Eng, 25 (1998), 161-197.

[Tian et al., 2005] Tian, Q., Ma J., Liang J., Kwok, R. C.W., Liu, O.: "An organizational decision support system for effective R&D project selection"; Decision Support Systems, 39, 3 (2005), 403-413.

[UNE 166005, 2004] UNE 166005:2004.: "Guía de aplicación de la Norma UNE 166002:2002 EX al sector de bienes de equipo"; (2004).

[UNE 166001, 2006] UNE 166001:2006.: "Gestión de la I+D+i: Requisitos de un proyecto de I+D+i"; (2006).

[UNE 166002, 2006] UNE 166002:2006.: "R&D&I Management: R&D&I Management System requirements"; (2006).

[UNE 166001, 2006] UNE 166006:2006.: "Gestión de la I+D+i: Sistema de Vigilancia Tecnológica"; (2006).

[Valencia-García et al., 2008] Valencia-García, R., Fernández-Breis, J.T., Ruiz-Martínez, J.M., García-Sánchez, F., Martínez-Béjar, R.: "A knowledge acquisition methodology to ontology construction for information retrieval from medical documents"; Expert Systems: The Knowledge Engineering Journal, 25, 3 (2008), 314-334.

[Valencia-García et al., 2011] Valencia-García, R., García-Sánchez, F., Castellanos-Nieves, D. Fernández-Breis, J.T.: "OWLPath: an OWL ontology-guided query editor"; IEEE Transactions on Systems, Man, Cybernetics: Part A, 41, 1 (2011), 121-136.

[Vidoni et al., 2011] Vidoni, R., García-Sánchezb, F., Gasparetto, A and Martínez-Béjar, R.: "An intelligent framework to manage robotic autonomous agents"; Expert Systems with Applications, 38, 6 (2011): 7430-7439.

[Zhu et al., 2010] Zhu, J., Huang, X., Song, D. and Rüger S.: "Integrating multiple document features in language models for expert finding"; Knowledge and Information Systems 23 (2010) 29–54.