Generating innovation with semantically enabled TasLab portal

Pavel Shvaiko[1], Alessandro Oltramari[2], Roberta Cuel[3], Davide Pozza[4], and Giuseppe Angelini[5]

[1] TasLab, Informatica Trentina S.p.A., Trento, Italy
pavel.shvaiko@infotn.it
[2] ISTC-CNR, Trento, Italy
oltramari@loa-cnr.it
[3] DISA, University of Trento, Trento, Italy
roberta.cuel@economia.unitn.it
[4] OMNYS srl - Information Technology, Vicenza, Italy
davide.pozza@omnys.it
[5] WEBSS srl - Web Software Solutions, Trento, Italy
gangelini@webss.it

Abstract. In this paper we present a concrete case study in which semantic technology has been used to enable a territorial innovation. Firstly, we describe a scenario of the ICT regional demand in Trentino, Italy; where the main idea of territorial innovation is based on the so-called *innovation tripole*. Specifically, we believe that innovation arises as a result of the synergic coordination and technology transfer among three main innovation stakeholders: (i) final users, bringing domain knowledge, (ii) enterprises and SMEs, bringing knowledge of the market, and (iii) research centers, bringing the latest research results. The tripole is instantiated/generated for innovation projects, and, technically, can be viewed as a competence search (based on metadata) among the key innovation stakeholders for those projects. Secondly, we discuss the implementation of the tripole generation within the TasLab portal, including the use of domain ontologies and thesauri (e.g., Eurovoc), indexing and semantic search techniques we have employed. Finally, we provide a discussion on empirical and strategic evaluation of our solution, the results of which are encouraging.

1 Introduction

The technology transfer process can be described as the economic and/or organizational process, of technology, knowledge and products exchange among various stakeholders. It leads to the transfer of knowledge, usually from research centers to industries, in order to conduct firms to the development and commercialization of end user applications and products [11, 13, 17]. In the last decades we have assisted to a radical change in the forms of alliances, processes and governance models of technology transfer. Specifically, let us consider the shift from traditional joint ventures and in-house R&D laboratories to a more flexible forms and more open environments, such as science innovation parks, incubators, and Living Labs[1], see also [3, 14]. This change introduced a new

[1] http://www.openlivinglabs.eu/

paradigm, called *open innovation* [2], through which many actors use both internal and external assets to advance their technology of production and service providing. In fact, several companies nowadays operate as open innovation intermediaries and some others report on using multiple instruments to open up their innovation systems, e.g., by using venture capital funds, managing collaborations with universities [16].

Such an open environment is the motivation for our case study: *Trentino as a Lab* (TasLab), which is considered as a governance model for technology transfer supported by the Autonomous Province of Trento, Italy [9]. TasLab is meant to play a mediating role among business, research and public sector, thereby enhancing innovation. The activities of TasLab can be generally articulated as follows:

- *conception:* namely, creation of shared knowledge, economic/organizational conditions and opportunities to stimulate creativity and innovative products or services. This means creating and maintaining a network of companies, research departments, and final users (both citizens and public administrations), scouting the technology, disseminating results and opportunities, etc.
- *development:* once an innovative proposal has been identified, a project should be formalized and managed. This means providing services, e.g., project management, networking for new partnerships, support for intellectual property rights (IPR), etc.
- *production:* namely, refinement of a product or a service in order to transfer it to the actual exploitation by the final users.

In order to support these activities, among others, various technologies, such as knowledge portals have been adopted. Within such portals, various stakeholders should have a role and their special skills and knowledge should be managed and integrated to develop innovative ideas. Notice that Trentino is characterized by a myriad of small and medium enterprises (SMEs) which develop their own business assets, processes and knowledge that might enable innovation. Hence, a territorial innovation portal should provide knowledge management facilities in order to match these assets, interests, and competencies of all the various stakeholders. It should support flows of knowledge involving links and staff flows between firms, research centers, and public institutions. This means that this heterogeneity/diversity (e.g., in goals, languages, terminologies of the various innovation stakeholders) should be treated as a feature that eventually might generate innovation, and hence, should be respectively handled [8]. In turn, this requires an adequate semantic heterogeneity management that we believe can be implemented with the help of the semantic technology.

The contributions of the paper include: (i) description of the concrete problem, namely innovation generation, in a specific application domain, such as eGovernment, where the semantic technology is of help; (ii) an implementatinon of the solution devised within the TasLab portal; and (iii) a discussion of the empirical and economic evaluation of the solution employed.

The rest of the paper is organized as follows. Section 2 provides the problem statement. Section 3 overviews the solution proposed. Section 4 describes the use of the semantic technology within our solution as well as some of its implementation details. Section 5 discusses an empirical and strategic evaluation of the solution employed. Finally, Section 6 reports on the major findings of the paper and overviews future work.

2 The problem: how to generate innovation

Our application domain is *eGovernment*. By eGovernment we mean here an area of application for information and communication technologies to modernize public administration by optimizing the work of various public institutions and by providing citizens and businesses with better (e.g., more efficient) services as well as with the new services (that did not exit before), see [20] for the latest developments in the area. More specifically, our goal is to introduce in a systematic manner (technological) innovation in the eGovernment projects being conducted in the Autonomous Province of Trento, Italy. We believe that semantic technology is among enabling (trasversal) technologies that can help achieving this goal.

2.1 Motivating example

Let us briefly discuss a typical situation occurring at a public administration. A final user, such as a department of public administration (e.g., Urban Planning) which among others, has the goal of sharing of spatial information has to follow the INSPIRE[2] directive, see also [19]. In order to achieve this, one of the key components of the INSPIRE architecture is a discovery service, that ought to be implemented by means of the Catalogue Service for the Web (CSW)[3]. A usual request of a typical final user here would be to have a geo-catalog. An advanced, or the so-called *lead user* would ask for a semantic geo-catalog. Although, there have been provided several implementations for the CSW-based geo-catalog, at present there is no reference implementation for a semantic geo-catalog. Hence, semantic geo-catalog is an example of an innovative project (at the world level) the final user may want to run.

Now, the question is what are the competencies available to run such a project. There is a need for competencies on semantic layer of the system from a research center. For example, there could be already a prototype for a discover service implemented at a research center to be further adapted to a particular case at hand. There is a need for an industrial company to provide a robust implementation for the semantic geo-catalog system-to-be, see for details [18].

Moreover, there can be several variations of the scenario discussed above, including:

- a user may want to identify several research centers in order to contact them concerning the state of the art with geo-catalogs and actual feasibility of the eventual innovative requests;
- a public institution may need to perform a market verification before launching a public tender, by asking some industrial companies an estimate for the realization of certain functionalities of a geo-catalog;
- an industrial company may have a promising platform and may look for its further extensions, and hence, may want to identify the relevant research centers, or final users that may help with this extension, e.g., either by including some new functionalities (research centers) or by customizing some modules of that platform to particular user needs, what ultimately can be re-proposed in other projects;

[2] INSPIRE - INfrastructure for SPatial InfoRmation in Europe: `http://www.ec-gis.org/inspire/`

[3] `http://www.opengeospatial.org/standards/cat`

– research centers or universities may have outstanding prototypes, and hence, may want to bring them down to production and final exploitation, or just find partners (e.g., industrial companies or final users) for a new project proposal.

From a technical viewpoint, the scenario mentioned above as well as its variations can be implemented as a search for competencies across the key innovation actors. Note that many research projects that employ ICT (e.g., semantic technology) correctly identify an application in which prototypes they develop can be eventually exploited. However, it is far rarely the case that *final users* are directly involved in the definition of requirements and use cases instantiating the applications under consideration within those projects. This is so because research projects are not usually concerned with bringing the original ideas developed within them down to the actual exploitation of these by the (expected) final users. Also enterprises that are often involved in larger research and development projects, for instance, of 4 years duration and with about 1K man-month effort, are primarily interested in acquiring know-how to be later exploited in their internal subsequent projects. Hence, in order to foster an early practical exploitation of the research prototypes, it is necessary to directly involve final users in the research and development cycles.

2.2 Problem statement

Involving final users into the research and development cycles requires addressing a *social challenge* of integrating relevant actors and facilitating the cross-fertilization among research centers, technology providers and user institutions, see Figure 1; this is exactly what is pursued by TasLab.

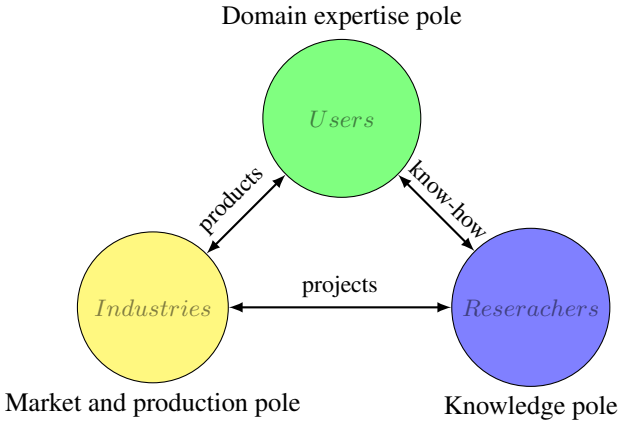

Fig. 1: The tripole model.

As Figure 1 indicates, the *tripole* model for innovation includes three key poles, namely: (i) users (U) that possess an application domain expertise, (ii) research centers and universities (R) that possess knowledge of the state of the art technologies and innovative solutions, and (iii) industries (D) that possess market and production knowledge. The actors of these poles are put together through the innovation projects, thus

being an enabling factor for amplifying potential gains by building on top of a larger volume of knowledge and experience compared with the other approaches, where final users usually have a limited role.

From a technical viewpoint, the problem of generating innovation is reduced to a competence search among the key innovation stakeholders and is as follows. Given:

- an innovation project theme (T), described as a phrase in a natural language, such as *semantic geo-catalog*;
- a finite set of user institutions $U = \{i \,|1, \ldots, N_U\}$, e.g., urban planning, each of which annotated with respective metadata $M_U = \langle name, competencies, \ldots \rangle$, all together denoted as $\langle U, M_U \rangle$;
- a finite set of research centers and universities $R = \{j \,|1, \ldots, N_R\}$, e.g., CNR (Italian National Research Council), each of which annotated with respective metadata $M_R = \langle name, recent\ projects, \ldots \rangle$, all together denoted as $\langle R, M_R \rangle$;
- a finite set of industrial companies $D = \{k \,|1, \ldots, N_D\}$, e.g., WEBSS, OMNYS, each of which annotated with respective metadata $M_D = \langle name, market, \ldots \rangle$, all together denoted as $\langle D, M_D \rangle$;

discover (based on the available metadata) the best match between T and $\langle U, R, D \rangle$; or in other words generate a tripole for a given project theme.

N_U corresponds approximately to 350 entities, including various public administration departments and municipalities that are respectively structured, and which ultimately offer services to about 500.000 inhabitants of Trentino; N_R = 13, grouping about 750 researchers and N_D = 684 for the ICT sector (we started from), grouping about 3.400 professionals [1]. The system is not closed, and thus, can be extended, e.g., by adding the research centers and the industrial ICT companies outside Trentino. However, the system bootstrap is done only with the data of Trentino. Finally, notice that the metadata inserted into the system is not bounded by any vocabulary or predefined lists of terms, thus, allowing users to use their own terminology which they believe describes best their activities. This preserves the diversity, being the key feature, of the innovation stakeholders, and thus, facilitates innovation generation.

3 The TasLab portal

We have implemented the tripole generation functionality within the TasLab portal[4], which is devoted to foster innovation by creating the conditions for a successful integration of the innovation stakeholders in the Trentino territory. The portal is available in English and Italian. It possesses various standard functionalities with public and private parts, including vision and mission of the initiative, related materials, news and events, services offered and so on. From now on, we concentrate only on a part of the back-end of the portal, which is related to the tripole generation.

Technically, the TasLab portal aims not only to model the pre-existing knowledge of the stakeholders but also to foster collaboration within public administration, companies and research centers, producing new structured knowledge. Specifically, it includes the following characteristics:

[4] `http://www.taslab.eu/`

– federation of lightweight ontologies (for the ICT domain) based on the sharing of a common reference model, i.e., Descriptive Ontology for Linguistic and Cognitive Engineering - DOLCE [6];
– multilingual tools (computational ontologies and human-language technologies that converge in the task of providing the semantic description of contents);
– query answering system for semi-automatic construction of lightweight ontologies (as a guide for non-expert users in the task of ontology building and population; not covered here and retained as future work).

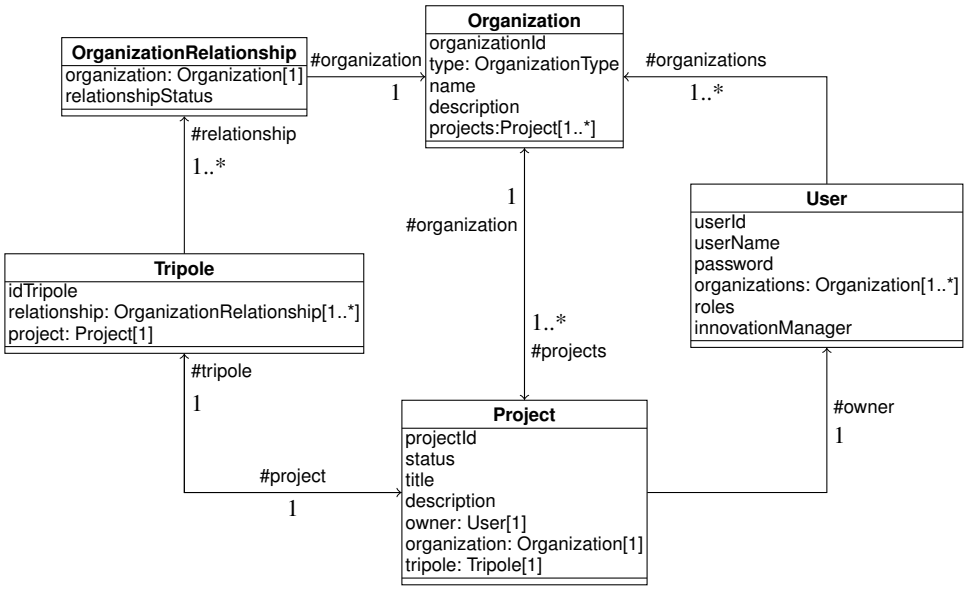

Fig. 2: Key business entities of the Taslab portal.

The main business entities involved in the Taslab portal are shown in Figure 2 and are as follows:

Organization, which can be of the following types: *company*, *public company* (a final user in our case) and *research center*. An organization includes one or more users; it can own a project (if one of its users created it) or collaborate on a project created by another organization.

User, which identifies a person who accesses the reserved area of the portal and uses the available functions (e.g., tripole generation). A user can assume the following roles: *portal admin*, who manages all the portal areas; *organization owner*, who owns an organization which has been approved by the portal admin; *organization member*, who is a member of an organization, typically an employee; *simple user*, who simply created an account on the portal, without becoming a member or an owner of any organization.

Project, which can be created only by a user belonging at least to an organization which is directly associated to it. Once a project has been created it can assume one of the following statuses: draft, started, in process, closed, suspended. During the

startup phase, the owner of the project is able to invite other organizations to participate in the innovative project proposal. This is enabled with the *tripole generation* functionality.

Tripole, which represents the group of organizations that have been involved in the project proposal through the following process: (i) the project owner has used the tripole generation functionality in order to semi-automatically identify the best matching organizations; (ii) the project owner has directly invited the identified organizations to participate in the project proposal; (iii) the invited organizations have accepted to be involved into the project proposal development.

4 The use of the semantic technology in the portal

In this section we describe how the semantic technology has been employed within the TasLab portal. First we present the portal architecture (§4.1), then we discuss the ontology of the portal (§4.2), indexing (§4.3) and semantic search (§4.4) techniques as well as finally the key interactions with the semantic search engine (§4.5).

4.1 The portal architecture

The TasLab portal has been implemented on top of the LifeRay[5] (v5.2.3 Community Edition) portal server. The tripole generation functionality has been implemented in Java. The system architecture, which follows three-tier paradigm, is shown in Figure 3.

Front-end tier. This is the front-end level of the application (also known as the presentation layer). It is responsible for serving the web contents to the Internet users. The software modules which provide the user interface and interact with the Middle tier are called *portlets* and follow a standardized development strategy - JSR 168/268[6].

Middle end tier. This layer (also known as the business logic layer) is responsible for performing the required business processes and returning the required results.

Back-end tier. This tier consists of all the Data Access Objects (DAO). These objects are responsible for storing and retrieving data from the data providers involved in the scenario. This tier keeps data neutral and independent from the application servers or business logic.

Finally, the Solr search engine is hosted on a separate servlet engine instance in order to decouple it from the portal server and to make it more scalable independently from any other module. Its pipeline has been extended by adding the required plug-ins, which are responsible for automatically performing, e.g., all the required semantic enrichments to the indexed contents (coming from the portal) and so on.

4.2 The ontology of the portal

The semantic layer of the portal is based on a hybrid model that integrates *Eurovoc* multilingual thesaurus[7] with a domain specific ontology. Eurovoc covers several domains

[5] http://www.liferay.com/web/guest/home

[6] http://developers.sun.com/portalserver/reference/techart/jsr168/

[7] http://europa.eu/eurovoc/

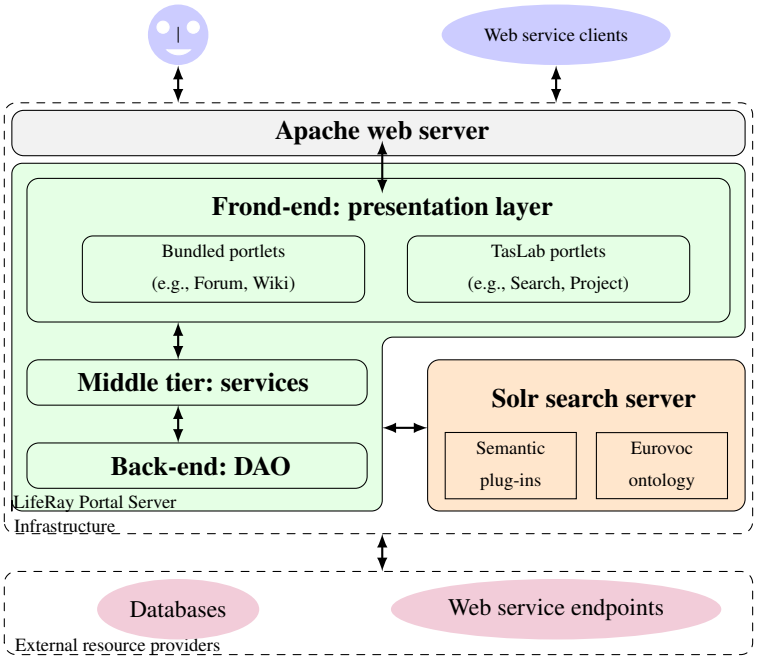

Fig. 3: The TasLab portal architecture.

and is available in 22 languages. It comprehends 21 fields, 127 micro-thesauri and 6645 descriptors (519 are the core terms). Moreover, there are 6669 occurrences of hierarchical links (broader term - BT/narrower term - NT) and 3636 associative relations (related term - RT), see Figure 4 for an example.

One of the main features of Eurovoc is the cross-lingual representation of terms and relations, allowing for coherent mappings between equivalent contents. Eurovoc has been preferred to alternative options, such as:

- a new thesaurus (or ontology) built from scratch: according to time and work constraints of the TasLab project, this direction has been considered too onerous.
- *eClass*[8]: a multilingual resource for product and service classification. Adapting this resource to the TasLab domain has not been considered as a feasible strategy, since the topic overlap between these is low.
- *EuroWordNet*[9]: This option has been discarded mainly for its genericity, where TasLab conversely needs coverage for specialized domains.

The ontology underlying TasLab portal concerns two layers, namely top level and domain level. In particular, at the top level, Eurovoc most general categories are aligned with DOLCE [6]. At the domain level, all the concepts used for describing TasLab entities (related to the institutional, research and administrative domains) are to be integrated into the Eurovoc structure (this is a future line work). In this sense, the ontology extends and enriches Eurovoc conceptual framework, both at top-down and bottom-up

[8] http://www.eclass-online.com/
[9] http://www.illc.uva.nl/EuroWordNet/

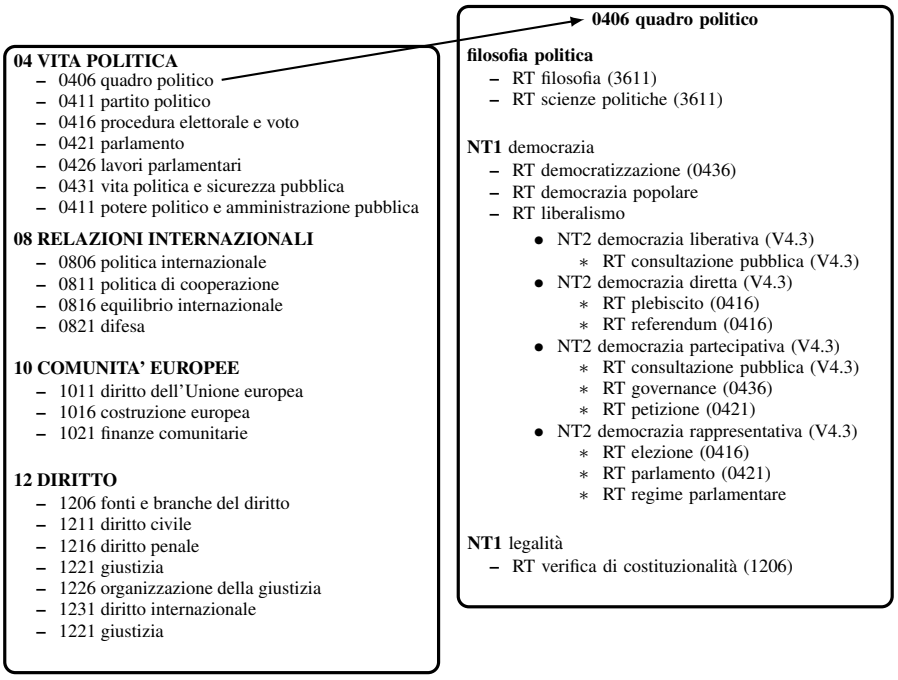

Fig. 4: A fragment of Eurovoc.

directions. The overall knowledge base has been implemented according to the SKOS W3C standard[10] (hence, converting Eurovoc from its native format to SKOS, etc.), which provides a unified method of knowledge representation, easier than the other languages, e.g., OWL, although exportable in that format.

4.3 Indexing

Every concept defined in the portal is assigned to a specific documental class and, afterwards, is converted into a Solr[11] document for indexing. In the initial stage, the specific fields of the documents are populated on the basis of their specific documental class. The following process, as shown in Figure 5, is applied to the relevant textual information/documents:

– Semantic enrichment, namely population of the TasLab domain specific ontology and its integration with Eurovoc.
– Stemming, namely reduction of the inflected (or sometimes derived) words to their stem, base or root form. For example, *progetto di sviluppo* being stemmed becomes *progett svilupp*. Since this process is language-dependent, there should ideally be as many stemming processes as the number of languages supported by the portal. In our case we have used it only for Italian and English, what improves the recall

[10] http://www.w3.org/2004/02/skos/
[11] http://lucene.apache.org/solr/

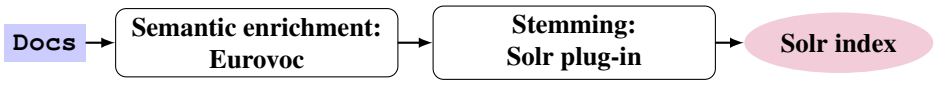

Fig. 5: Indexing.

of the system, since queries are normalized and adapted to the linguistic knowledge units of the portal.
– Data entry, namely documents are indexed on Solr and immediately made available for querying.

4.4 Semantic search

This functionality enables the tripole generation by executing semantic search for the required competencies. Search queries are processed as follows, see Figure 6:

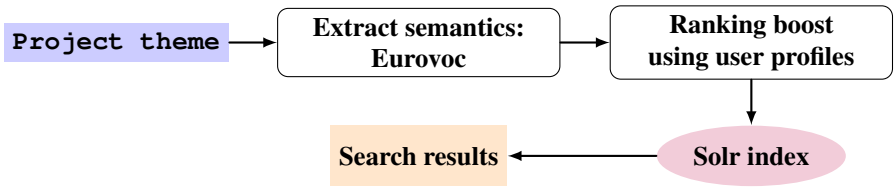

Fig. 6: Semantic search.

– Semantic extraction, namely matching between every searched term (undergo stemming) and the following items of the system ontology: domains, micro-thesauri, descriptors, broader terms, non descriptors, related terms and the related query-expansion, e.g., for the query *office automation*, the results here include: *office automation* (in exact and stemmed way) or Eurovoc sector *education and communications* or Eurovoc micro-thesauri *information technology and data processing* or Eurovoc descriptor *office automation* or broader term *computer applications*.
– Matching on the index, the expanded search (like in the example above) for each language (English and Italian) is submitted to Solr; results are given on the basis of textual and Eurovoc code matching.
– Ranking boost, results are contextually organized according to user profiles, specifically on the one side, by considering the empirically established weight assigned to the Eurovoc hierarchy (sector, thesauri, descriptor) and, on the other side, by exploiting the *interests* metdata entries of the user profile.
– Faceting, for every search, Solr also provides clustering of results according to categories/facets. These categories are dynamic, as these are processed by Solr for every query.

The metadata used for navigation includes: domain, micro-thesaurus and Eurovoc descriptor, release date (organized per period); language; documental class; document format. Finally, it is possible to refine search results on the basis of facets - contextually visualized at the result page as well as by providing access to similar documents (through the TF/IDF measure - term frequency/inverse document frequency), exploiting existing links.

4.5 Interactions with the semantic search engine

Let us discuss the key interactions between the specific functions supported by the portal and the semantic engine implemented as a Solr plug-in:

Semantic search over the contents of the portal. All the contents of the TasLab portal, suitably loaded through the LifeRay web content standard portlet, are retrieved using Eurovoc-based semantic search algorithms (see §4.4). Moreover, the Solr index is extended in order to manage the business entities discussed in Section 3.

Tripole generation. The metadata associated to the project is added to Solr and is exploited together with a list of keywords that the user can optionally indicate to guide the tripole generation according to the *best-each* modality for selecting the companies, research centers, public administrations. Specifically, the algorithm evaluates metadata of each organization entity and metadata of the related users, then it compares these with the project metadata and provides an ordered list (by ranking) for each type of the organization.

New user registration. When registering, the user can specify personal interests on the basis of Eurovoc. This information constitutes the user profile. The overall metadata and the URI's of the organization to which the user belongs are indexed by Solr. Notice that every change here forces the global re-indexing.

Adding new content. This process contextually requires necessary information for indexing, including: documental class, URI (automatically assigned by the system; this unique key is used for retrieving the document in the original format from the portal CMS repository), language (based on which the system determines the kind of stemming process to apply), title, release date (automatically assigned by the system). Every possible modification to metadata or contents forces the global re-indexing.

Adding new organization or project. The process of registration of a new organization foresees, after a first stage of moderation by the administrator, adding descriptive metadata to the Solr index, so that the new organization is made available for the semantic search and the tripole generation functionality. Every change here forces the global re-indexing. Similar argument applies when a new project is added.

5 Discussion

In this section we provide a discussion on some empirical (§5.1) and strategic (§5.2) evaluation of the solution employed together with a related work overview (§5.3).

5.1 Empirical evaluation

We have run our tests on a PC with CPU Intel quadcore 3 Ghz, 8GB of RAM. TasLab semantic search module underlying the tripole generation functionality builds on top of the performance of "Solr as is" since for each indexing/search, it executes about 20.000 to 25.000 queries for each language (in our case for English and Italian). The overhead here depends on: number of the documents on index; size of the index; search

parameters, such as query complexity, numbers of request fields, etc. In our tests we had 100.000 random documents on index, 1.2GB index size, queries of various length, such as uri, publishing date, title, facets on Eurovoc fields. Based on these tests the overhead was only about 100-200 ms, what suggests for the scalability of our solution.

The solution employed represents an improvement over the canonical (and commonly known) exact string matching search methods available in LifeRay. The user has the possibility of posing a query by using a natural language format and the semantic search algorithms used are able to return the most relevant documents related to the request. In fact, in all the examples we run so far our solution had demonstrated better characteristics (e.g., precision, recall) with respect to conventional approaches, such as a string-based search of LifeRay. For example, a user searching for *nuclear industry* obtains also results/documents on *plutonium* or *uranium* (the terms related to the nuclear industry argument) even if their contents never contain the searched string.

5.2 The SWOT analysis

Usually an assessment of an application implies conducting an ex-post analysis, which takes into consideration the effectiveness of the technology based on the daily activity of users. At the moment of writing, the TasLab portal is still under testing, therefore a *prospective* assessment method has been adopted: the *strengths*, *weaknesses*, *opportunities*, *threats* (SWOT) analysis. Typically, the SWOT analysis is used to review the market position, foreseen directions of a business idea. We decided to use it here in order to assess what strategic actions should be adopted in order to make the TasLab portal a success. To this end, we start by enlisting the needs of the main innovation stakeholders, which are as follows:

- *Public administrations,* which aim at: catalyzing the innovation and demanding for innovative products and services for the local government, monitoring and funding innovation programs, increasing the internationalization of local SMEs by fostering new partnerships.
- *Research centers,* which aim at: developing a more structured channel of technology transfer among academia and industry, testing innovative ideas in real-life settings, obtaining funds for future research activities.
- *Industries,* which aim at: obtaining new funding opportunities for the development of innovative solutions, enforcing connections with academia, suggesting/selling to the public administration specific (innovative) solutions, participating to the transformation of innovation into business value.

Following the strengths/weaknesses of ICT itself and market opportunities/threats of the TasLab portal, below we discuss how the above mentioned aims might be met.

Internal strengths. By using the semantic technology the portal improves quality of stakeholders' content management and information access. For example, the semantic search engine improves the accuracy of tripole generation (over the conventional methods, see §5.1) and in general of information discovery on the portal. Moreover, TasLab is strongly related to the territoriality and the inner characteristics of the organizations of the innovation network, what, e.g., means that the project can count on the availability of (local) domain experts to tune the application.

Internal weaknesses. Companies, research centers, and final users might anyhow eventually desert the portal, perceiving it as too complex to understand, due to bureaucratic procedures or being conflicting with their interests. In fact, a combination of motivation/incentive mechanisms is still in a very early development stage here. In order to build, test and provide to the market an effective territorial innovation system, several more years are needed; e.g., effective IPR policies should be adopted. In our view, the weaknesses reside on the organizational side, rather than on the technological one.

External opportunities. Companies, research institutions and local governments spend billions of euros each year on data integration and portals in order to foster and sustain innovation. Thanks to the availability of public financing of research and innovation in Trentino and thanks to the presence of top notch researchers, local companies are becoming also more competitive on the global market. We expect an increasing demand here coming from the diffusion of the portal in the other areas of Italy and in Europe.

External threats. In general, the complexity of creating semantically enabled applications and the sometimes unclear business value of these in comparison with more traditional approaches to master data and metadata somewhat limited the semantic technology potential and appeal [7]. This should be overcame with concrete benefits coming from the actual portal usage.

Through an effective knowledge management within the portal, companies and research centers should become aware of new funding opportunities, develop new profitable partnerships, ultimately leading to the development of new products and services. In turn, public administrations can monitor the results of innovative projects, direct the innovation activities into specific priority areas, etc. The local public administration is interested in collaborative solutions and in TasLab as one of the important innovative initiatives in the region. A strategic action to be taken includes designing a set of incentives in order to guarantee sustainability and high usage of the portal.

5.3 Related work

In the last years, semantic technology has played a key role in the ICT area. If semantics undoubtedly became a buzzword in the most advanced contexts of communication technologies (e.g., eGovernment, eHealth), only few systems are actually able to provide effective semantic functionalities. For example, looking at web-based communities in the area of knowledge management, it is not easy to find full-operational semantic portals, as also demonstrated in a recent survey in [12]. Here, the focus is on the general features of some of the most popular semantic portals: Esperonto[12], OntoWeb[13] (both developed for research tasks), Empolis K42[14], and Mondeca ITM[15] (commerce-oriented). These portals provide only static ontologies without any sufficient versioning mechanisms. In fact, none of these supply semantic web services, which conversely should be a core function of such systems. Within these four cases, only Ontoweb includes a hybrid semantic resource integrating an ontology with a suitable thesaurus,

[12] http://www.esperonto.net/
[13] http://ontoweb-lt.dfki.de/
[14] http://www.empolis.com/
[15] http://www.mondeca.com/

though making information access difficult. At this level, the navigation tools rely on conceptual models (the underlying ontologies) with no support for linguistic access. In general, the interface between ontologies and lexical resources has also an impact on semantic search, as recently stated in [10], since it represents a novel approach to user-centred information access based on computation of semantic relations between lexicalized concepts and not on syntactic parsing. In the variegate context of semantic search tools, such (a) user-based approaches to retrieve information and knowledge (e.g., SemSearch[16]) complement (b) user-centred search to retrieve ontologies (e.g., Swoogle[17] and Watson[18]) and (c) search based on structural query languages[19].

As from [12], the most relevant problem of the above mentioned semantic portals concerns the total lack of community-based features, corresponding to the absence of suitable dynamic functionalities for the evolution of portal contents according to users' practice and knowledge. In this sense, as we have illustrated in the previous sections, TasLab can be considered as a good candidate to overcome that limit, since it *delegates* to the tripole actors the construction of the information framework for the semantic portal[20]. This remark is in line with thesis of the work in [5] too: to globally take into account semantic technology into the eGovernment applications; i.e., it is not enough to focus on web services, but an ontological focus on people and organizations is also needed. For this reason, we think that - besides the critical considerations in [12] - what is important for the efficiency of semantic systems is to focus on the human dimension of information, reminding that the Web is not only an interlinked cluster of machines, but rather a network of humans negotiating linguistic meanings through machines. One of the key issues in current R&D in semantic technology is to provide integrated solutions to address two core requirements of the Web: (i) natural language semantic facilities and (ii) flexible machine-encoding and processing of concepts. We can conclude that the above mentioned features of TasLab are going to satisfy these requirements.

6 Conclusions

In this paper we have discussed our solution to innovation generation in public administrations. Specifically, we reformulated the problem of innovation generation as competence search or the so-called tripole generation among the key innovation actors, namely: final users, industries and research centers. This is enabled with the semantic technology, in particular with the use of ontologies, indexing and semantic search. The evaluation results, though preliminary, demonstrate the strengths of our approach as from the technical viewpoint (e.g., scalability) as well as from the SWOT analysis.

Future work includes at least: (i) an in-depth evaluation of the techniques employed as well as including some other matching methods from [4] in order to fine-tune the solution; (ii) an enrichment of Eurovoc by customizing TMEO, a recently developed tutoring methodology for the enrichment of ontologies [15], to the TasLab requirements;

[16] `http://code.google.com/p/semsearch/`

[17] `http://swoogle.umbc.edu/`

[18] `http://watson.kmi.open.ac.uk/WatsonWUI/`

[19] Such as SPARQL and ARQ (see, e.g., `http://jena.sourceforge.net/ARQ/`).

[20] Only providing a top-down resource (DOLCE-based domain ontologies + Eurovoc thesaurus) at the initial stage.

and (iii) the development of an organizational structure which with an appropriate set of incentives should guarantee sustainability and high usage of the portal.

Acknowledgments. This work has been supported by the *TasLab network* project funded by the European Social Fund under the act n. 1637 (30.06.2008) of the Autonomous Province of Trento. We are thankful to the TasLab group members: Isabella Bressan, Ivan Pilati, Luca Mion, Valentina Ferrari, Marco Combetto and to Fausto Giunchiglia, Francesca Gleria, Italo Della Noce, Mauro Piffer, Andrea Simioni for many fruitful discussions on the various innovation aspects covered in this paper as well as to Gabriele Frati for some technical support.

References

1. P. F. Camussone, I. Dalmonego, and E. Zaninotto. *Le tecnologie digitali nell'economia del Trentino*. EGEA S.p.A., Italy, 2007.
2. H. W. Chesbrough. *Open Innovation: The New Imperative for Creating and Profiting from Technology*. Harvard Business School Press, 2003.
3. M. Eriksson, V.-P. Niitamo, S. Kulkki, and K. A. Hribernik. Living labs as a multi-contextual R&D methodology. In *Proc. of ICE*, 2005.
4. J. Euzenat and P. Shvaiko. *Ontology matching*. Springer, 2007.
5. R. Ferrario and N. Guarino. A new ontological perspective for social services. In *Proc. of MeTTeG*, pages 41–51, 2008.
6. A. Gangemi, N. Guarino, C. Masolo, and A. Oltramari. Sweetening WordNet with DOLCE. *AI Magazine*, 24(3):13–24, 2003.
7. Gartner. Hype cycle for web and user interaction technologies, 2009.
8. F. Giunchiglia. Managing diversity in knowledge. *Keynote talk at ECAI*, 2006.
9. F. Giunchiglia. Il ruolo degli enti di ricerca per lo sviluppo dell'ict del trentino. In *Le tecnologie digitali nell'economia del Trentino*, 2007.
10. F. Giunchiglia, U. Kharkevich, and I. Zaihrayeu. Concept search. In *Proc. of ESWC*, pages 429–444, 2009.
11. J. Howells. Intermediation and the role of intermediaries in innovation. *Research Policy*, 35(5):715–728, 2006.
12. H. Lausen, Y. Ding, M. Stollberg, D. Fensel, R. Lara Hernandez, and S.-K. Han. Semantic web portals: state-of-the-art survey. *Journal of Knowledge Management*, 9(5):40–49, 2005.
13. M. Levin. Technology transfer as a learning and development process: an analysis of Norwegian programme on technology transfer. *Technovation*, 13(8):497–518, 1993.
14. V.-P. Niitamo, S. Kulkki, M. Eriksson, and K. A. Hribernik. State-of-the-art and good practice in the field of living labs. In *Proc. of ICE*, 2005.
15. A. Oltramari. TMEO - tutoring methodology for the enrichment of ontologies. In *Proc. of LREC*, 2010 (to appear).
16. R. Rohrbeck, K. Hölzle, and H. G. Gemünden. Opening up for competitive advantage: How Deutsche Telekom creates an open innovation ecosystem. *R&D Management*, 39(4):420–430, 2009.
17. P. Shapira. Modernizing small manufacturers in Japan: The role of local public technology centers. *Journal of Technology Transfer*, 17(1):40–57, 1992.
18. P. Shvaiko, L. Vaccari, and G. Trecarichi. Semantic geo-catalog: a scenario and requirements. In *Proc. of the workshop on Ontology Matching collocated with ISWC*, pages 240–241, 2009.
19. L. Vaccari, P. Shvaiko, and M. Marchese. A geo-service semantic integration in spatial data infrastructures. *Journal of Spatial Data Infrastructures Research*, 4:24–51, 2009.
20. M. Wimmer, H. J. Scholl, M. Janssen, and R. Traunmüller, editors. *Proc. of the 8th International Conference on Electronic Government (EGOV)*, 2009.