VRIJE UNIVERSITEIT BRUSSEL

FACULTEIT WETENSCHAPPEN

VAKGROEP INFORMATICA EN TOEGEPASTE INFORMATICA

SYSTEMS TECHNOLOGY AND APPLICATIONS RESEARCH LAB

STAR Lab Technical Report

Using a novel ORM-based ontology modelling method to build an experimental Innovation Router

Peter Spyns, Sven Van Acker, Marleen Wynants *, Mustafa Jarrar & Andriy Lisovoy.

affiliation:	* currently at VUB – R&D department – CrossTalks
keywords	innovation, ontology modelling
number	STAR-2004-07
date	6/04/2004
corresponding author	Peter Spyns
status	final
reference	

Pleinlaan 2, Gebouw G-10, B-1050 Brussel
Phone: +32-2-629.1237; Fax: + 32-2-629.3819
http://www.starlab.vub.ac.be

Using a novel ORM-based ontology modelling method to build an experimental Innovation Router

Peter Spyns, Sven Van Acker, Marleen Wynants, Mustafa Jarrar & Andriy Lisovoy

Vrije Universiteit Brussel – STAR Lab
Pleinlaan 2, Gebouw G-10, B-1050 Brussel, Belgium
{firstname.lastname}@vub.ac.be; http://www.starlab.vub.ac.be

Abstract. Specific tools help to increase the effectiveness of a shortened innovation cycle. The paper presents a web based tool for the creation of a scenario of what an innovation environment looks like, by enabling enriched queries and by allowing the identification of specific innovation gurus and key role institutions. The tool relies on an ontology-based knowledge representation that has been built using a recently adapted conceptual modelling methodology.

1 Introduction

Since the studies of Schumpeter[1] (e.g., [27]), (technological) innovation has been recognised as a crucial part of a company's assets. As a consequence, companies try to set up processes to manage and stimulate innovation activities: internal R&D teams define, select, develop, test and exploit innovative ideas that should lead to a next generation of services and products. In essence, innovation has to do with knowledge about technology, and business and production processes. Innovation is the result of a mutual influence, difficult to capture and model, between science, industry and the market. The government is another important actor who tries to stimulate innovation via its policies (e.g. regarding science, (higher) education, SME programs etc.).

In the economic rat-race, the "time to market" and life span of a product have drastically been reduced, implying that the span of an innovation cycle in its turn has to be shortened, while facing tighter budget constraints. Cumulative cash flow diagrams (see e.g., [34]) get "squeezed". Specific tools to manage the innovation cycle help to increase the effectiveness of a shortened innovation cycle. It would be essential to align the functionality of these supporting tools technologies with the needs of knowledge users in a variety of industries including business and professional services, engineering, information technology, manufacturing, health care, publishing, etc. [2].

---

[1] Actually, Karl Marx was the first classical economist who extensively studied innovation (Capital, Vol. 1, 1867, especially Chapter 13) [remark from an anonymous reviewer].

[2] The project maintains a web site at http://www.innovanet.eu.com where systematic innovation related information and resources may be found. The members of the consortium are: Inmark Estudios y Estrategias S.A. (Spain), BioVista (Greece), FhG – IPSI (Germany), VUB STAR

2 Objectives

The main goal of the "Innovanet" project (EU 5FP IST 2001-38422 [3]) was to prepare a strategic roadmap on the possible 'systematisation' of the process of innovation and scientific discovery. It aims at a better understanding of where and how creativity and innovation come into play and of the kinds of software environments and IT technologies that could promote their emergence in research and industrial environments. One of the tasks of the *VUB STAR Lab* team consisted of building an ontology-based knowledge representation to organise data related to innovation processes in a manageable fashion. In addition, a web-based tool (called Innovation Router [4]) should enable the creation of a scenario of what an innovation environment looks like, by enabling enriched queries. The Innovation Router has to allow the identification of specific innovation gurus and key-role institutions. An adaptation of an existing conceptual modelling methodology (developed for the Flemish IWT GBOU 2001 #010069 "OntoBasis" project [5]) [30] has been used to create the innovation ontology.

This paper is organised as follows: after an overview of the material (section 3) and a summary of the modelling methodology (section 4.1) used to create the innovation ontology (section 4.2), the results (section 5) are discussed. In particular, two use cases for the Innovation Router are presented, one on searching for competent staff (section 5.2.1) and one on looking for patents (section 5.2.2). Related work is compared in section 6. Plans for future work (section 7) are given before a conclusion (section 8) ends the paper.

3 Material

The Innovation Router structures the relevant material of FP4 and FP5 from the Cordis[6] database, European patent applications and scientific literature. The data provided within the framework of the Innovanet project was limited geographically, at the level of disciplines and technologies with regard to defined activities and their actors. The results of a bibliometric analysis performed on the raw data [7] were stored in separate databases according to the data's provenance. These databases were, depending on the respective sources, of different data quality, contained redundant information, and had incomplete descriptions of database attributes. All these databases have been grouped in a single MS Access database that contains:

Lab (Belgium), PIRA International UK (UK), ITC-IRST (Italy) and Bit Media e-Learning Solution (Austria).

[3] The Innovanet project was a roadmap project that lasted from 01/11/2002 until 31/11/2003.

[4] A name « Router » symbolises that innovation information is intelligently guided or « routed » to a (human) user. On purpose we avoided to call it a portal as a portal involves many more things – e.g. see [28].

[5] See http://wise.vub.ac.be/ontobasis for more information.

[6] http://www.cordis.lu

[7] The analysis has been done by partner Clemens Wildhalm (Bit Media e-Learning Solution).

– 403 scientific papers
– 2.212 European patent applications
– 849 FP4 projects
– 3130 FP5 projects
– 867 Keywords (only for articles) [8]
– 6374 Phrases (articles, patents, proposals) [9]
– 24 Research Areas (only for proposals)

4 Methods

Due the amount, variety and complexity of innovation processes and their related data, precise and unambiguous semantics of the data were lacking. In order to enhance the potentialities of unambiguous data exchange and future exploitation activities, "an innovation ontology" was created. In current computer science parlance, an ontology [9], [10], [11] is understood as a vocabulary with semantically precise and formally defined terms that stand for concepts and their relationships of an application domain. VUB STAR Lab has adapted an ontology modelling methodology based on an existing conceptual schema modelling methodology [21] called Object Role Modelling (ORM [12]) – see section 4.1. The innovation application ontology has been created according to this methodology and served as a reference for the data model of the Innovation Router and as the underlying conceptual model of the web-based interface – see section 4.2.

4.1 Defining an Ontology Modelling Methodology

The ORM conceptual modelling methodology has been selected because of its strong foundation in natural language, which it inherited from its predecessor method called "a Natural Information Analysis Method" (NIAM) [35]. The latter was developed in the 1980s as a methodology to model databases. It introduced the distinction between a lexical (label) and non-lexical (thing) modelling objects, and supported subclassing and an extensive set of declarable integrity constraint types. While not an actual natural language-based approach with tools, NIAM did support the negotiation and agreement process to arrive at information requirement specifications. These took the form of conceptual semantic networks, or verbalisations of them, that were readable by non-computer experts and yet could be readily transformed into database designs [10]. Initially, ORM as such has been used before by the authors when creating ontologies (e.g., [16]), but the need occurred to have the methodology evolve into a "genuine" ontology modelling method. In this section, we summarise the results (section 4.1.1) and discuss some major differences with the original NIAM/ORM methods (sections 4.1.2 & 4.1.3).

[8] Unfortunately the data is rather of low quality.
[9] A phrase is a group of one to five words. Also here, the quality is rather low.
[10] We refer the interested reader to [12] and [35] for more details.

4.1.1 Modelling Steps

The seven basic steps of Halpin [12] were maintained but should be applied in a partially redefined way. An additional step (step 3: grounding the vocabulary) has been added. During this step, terms belonging to an application domain are explicitly attributed a meaning (be it still intuitively by a gloss and in principle irrespective of specific application requirements at hand) by creating a new concept with a corresponding unique definition for it or by associating the term to an existing concept and its definition. This is needed to reach an agreement on meaning prior to its sharing. This step is typical of and essential for ontologies.

The division of the steps into two parts corresponds to the double articulation of a DOGMA ontology, i.e. a clear separation between the lexon base and the commitment layer [29] [11]. The commitment layer "houses" the attribution of formal characteristics of use (constraints – e.g. cardinality) to a selection of concepts linked with terms (lexons from the ontology base) belonging to an external application (i.e. ontologically committing these terms).

1. Part I: conceptualise the domain
    1. verbalise information examples as elementary facts
    2. create the lexons (for a context and a language)
    3. ground the terms and roles constituting the lexons
2. Part II: add the constraints
    1. uniqueness
    2. mandatoriness
    3. subset, equality, exclusion and subtyping
    4. occurrence frequency and ring constraints
    5. final consistency checks

4.1.2 Modelling Constituents

According to Halpin, a conceptual schema of a database consists of three main constituents [12:p.31]:

- *basic fact type*s: the kinds of primitive sentences or facts
- *constraints*: the restrictions that apply to the fact types
- *derivation rules*: rules, functions or operators (including mathematical calculation or logical inference) to derive new facts from other facts.

Basic facts are asserted concerning an application domain. A (binary) fact states that a specific object is related in a particular way (plays a specific role) with another object. Objects can be entities or values. A unary fact asserts that an object plays one role. The concrete factual information is represented in an "information template". It consists of a combination of entity types (in NIAM: non lexical object type or NoLOT), values types (in NIAM: lexical object type or LOT) and predicates. A predicate combining an entity type with a value type is called the reference mode (in

[11] Due to space restrictions, we cannot elaborate on DOGMA (Developing Ontology Guided Mediation for Agents). We refer the interested reader to the publication section of our website: http://www.starlab.vub.ac.be .

NIAM: bridge type). A predicate combining two (or more) entity types is called a fact type (in NIAM: idea type).

Translated in terms of the DOGMA initiative, it means that the *basic fact types* belong to the ontology base and the *constraints* belong to the commitment layer. The *derivation rules* are actually not considered as part of the actual ontology, in opposition to what other ontology researchers often claim. The derivation rules are situated in the application domain realm. Basically, inference rules use (e.g., as the signature of a first order language) the vocabulary as it has been defined and constrained in the ontology. Further research needs to be done on how to model these derivation rules.

4.1.3 Referencing

Referencing in an ORM conceptual data model happens by means of a *bridge type* (using the NIAM terminology) between a LOT (in ORM: entity type) and a NoLOT (in ORM: value type). E.g., a person is identified by his first name. The actual values (or strings) for the first names are stored in the database (e.g., table 'Person' with a column label 'firstname') [object level]. As ontologies, in principle, are not concerned with instances (=data) but with meta-data (concept labels), referencing can only be done when an application has adopted a commitment (via lexical mapping rules).

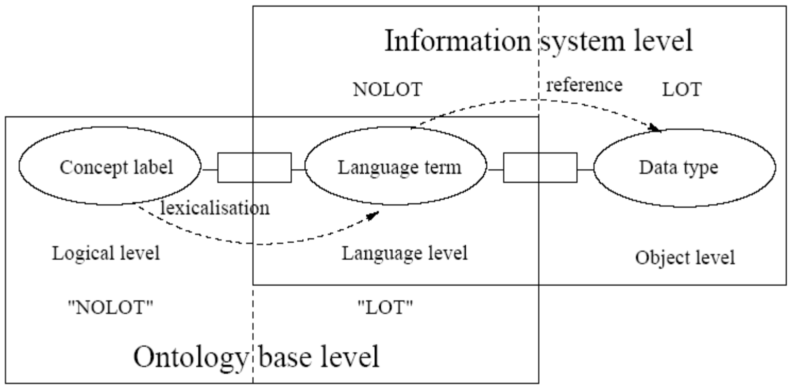

Fig. 1. three layer reference scheme (reproduced from [3])

Databases that use different terms for the same notion can share data if the meaning of the local database vocabulary (table and column labels) is mapped to the meaning of the corresponding term in the ontology vocabulary (the latter being precisely defined). A reference scheme (linking a notion to a data type – see Fig. 1) will now have three levels: a LOT that refers to a NoLOT (both belong to the conceptual data model of the information system) which in turn is linked to a centrally defined concept label – the latter two belonging to the ontology base level. At the time of modelling an ontology, the instance population is not always available. Reference schemas are thus no primary concern for an ontology modeller (but rather for an application developer). It implies that only NoLOTs can be used in an ontology (see Fig. 1). LOTs only appear in an additional application layer, especially when vocabulary of legacy systems are associated with existing domain concepts [3].

In case a database model has been designed on basis of an ontology, it would be natural to see that the ontology terms are used inside of the conceptual data model (cf. global as view). It means that two levels of the reference scheme (concept label and DB NoLOT) are collapsed and that the terms play a double role (the middle part of Fig. 1 disappears). The inverse scenario is to create an application ontology [11:p.10], by extracting an ontology from a conceptual data model and defining the semantics of its terms that are promoted to concept labels (cf. local as view – e.g., [19] & [32]). An application ontology can subsequently be merged with more general domain ontologies and/or other application ontologies.

Even in the ontology literature, authors do not always make a clear distinction between a global concept and a local conceptual database model term (the latter in many cases expressed in the local natural language). In particular, application ontologies are a dubious case: terms of the conceptual scheme of a database are often treated as concept labels but without an accompanying specification of their meaning (gloss or dictionary style definition), a shared and agreed meaning cannot be reached. A term on it own is not sufficient. And as for ease of reading and simplicity, many knowledge engineers label concepts by means of a (representative) natural language term the result being that the (global) conceptual and (local) language or application levels become quickly mixed up, which can be harmful when aligning and merging ontologies.

4.2 Creating an Application Ontology for Innovation

A data model is "tuned" towards a specific application, and therefore has less or no needs for explicit semantics (since sharing is not required). A conceptual data model is a "parsimonious" model, i.e., only distinctions relevant for that particular application are considered. An ontology is a "fat model" as it is to be shared across many applications, therefore needing a larger coverage and higher granularity. An application ontology, since it is directly derived from a conceptual model, is rather to be considered as a parsimonious model.

With an eye on timely realising the Innovation Router, we choose to create a slightly extended application ontology. Relevant properties and relationships of the entities were extracted from the input databases. Additional concepts and relationships were added to extend the original ontology mainly to enhance its genericity and re-usability.

In the next section, we will illustrate the first part of the new methodology or the conceptualisation of the domain. Adding the constraints is quite straightforward for knowledge engineers familiar with ORM/NIAM. The reason for not discussing the second part is due to space limitations. In addition, there is point in explaining this for an application ontology, as the semantic restrictions simply correspond with the restrictions on the conceptual schema. Upgrading a conceptual schema to an application ontology consists in providing the semantics for the labels used in the schema. And in this particular case, the schema was simple and the restrictions were rather basic.

- [step 1] On basis of the relational schema of the various databases, the entities, relationships and their properties have been verbalised: *people are involved in projects*, *a person covers a scientific topic*, *a person works for an organisation*, … [12] . These sentences can no longer be split into smaller units of information (i.e. elementary facts). They can be derived from tables like (expressed as relations) *Employs(Project, Person)* and *Expertise(Person, Skills, Institute)* respectively.

- [step 2] The elementary facts expressed in a restricted form of natural language are transformed into formally defined lexons that can be read from both sides:

  $$\mathbf{<(\gamma, \lambda): \text{head-term, role, co-role, tail-term}>} \qquad \mathbf{(1)}$$

  The λ is a label that indicates the language, and the context identifier, γ, groups lexons that are intuitively “related” in an intended conceptualisation of a domain. See [3] for details on the most recent formalisation of a lexon.
  e.g.,:
  <*(innovation, English-UK): person, works_on, involves, project*>
  <*(innovation, English-UK): person, works_for, employs, organisation*>
  <*(innovation, English-UK): person, has_expertise_in, is_covered_by, topic*>

- [step 3] The constituents of the lexons are associated with word sense definitions and concepts (potentially newly created) [13]. It is natural to use existing resources, such as WordNet [21]. Other potential resources are CYC [17], DOLCE [8], EuroWordNet [36], or UMLS [14] for the medical domain. WordNet contains synsets and definitions of what a term means. A disadvantage is that WordNet covers mostly non-technical vocabulary. Therefore, the modeller, with the collaboration of a terminologist or lexicographer, will have to come up with definitions for technical terms. We recommend doing it in the same style as WordNet [14]:

  <u>Terms:</u>
  - *person*: **person#1**
    individual, someone, somebody, mortal, human, soul -- (a human being; "there was too much for one person to do") [sense 1 of 3]]
  - *organisation*: **organisation#1**
    (a group of people who work together) [sense 1 of 3]

---

[12] Due to confidentiality agreements amongst the consortium partners, the entire ontology is currently still confidential. Only short excerpts will be provided for illustrative purposes.
[13] See Nirenburg [24] for a discussion on language neutrality vs. language independence of word senses.
[14] http://www.cogsci.princeton.edu/cgi-bin/webwn

– *topic*: **topic#2**
  topic, subject, issue, matter -- (some situation or event that is thought about; "he kept drifting off the topic"; "he had been thinking about the subject for several years"; "it is a matter for the police") [sense 2 of 2]
– *project*: **project#2**
  project, projection -- (a planned undertaking) [sense 2 of 2]
– *expertise*: **expertise#1**
  expertness, expertise -- (skillfulness by virtue of possessing special knowledge) [sense 1 of 1]

Roles:

– *works_for*: **work#2(v)**
  work, do work -- (be employed; "Is your husband working again?"; "My wife never worked"; "Do you want to work after the age of 60?"; "She never did any work because she inherited a lot of money"; "She works as a waitress to put herself through college") [sense 2 of 27 as a verb]
– *employs*: **employ#2(v)**
  hire, engage, employ -- (engage or hire for work; "They hired two new secretaries in the department"; "How many people has she employed?") [sense 2 of 2 as a verb]
– *covers*: **cover#5(v)**
  cover, treat, handle, plow, deal, address -- (deal with verbally or in some form of artistic expression; "This book deals with incest"; "The course covered all of Western Civilization"; "The new book treats the history of China") [sense 5 of 26 as a verb]
– *works_on*: **work_on#1**
  work at, work on -- (to exert effort in order to do, make, or perform something; "the child worked at the multiplication table until she had it down cold") [sense 1 of 2]
– *involves*: **involve#2**
  involve -- (engage as a participant; "Don't involve me in your family affairs!") [sense 2 of 7]

A lexon is thus an intermediary step towards a language independent conceptualisation of a domain. Note that no formal axiomatisation is done. In WordNet all entries are linked to an internal upper ontology of which the semantics are currently under revision [8]. Reasoning components (in our vision relegated outside the ontology [15]) can make use of these axioms.

Once the ontology was finalised, a new unified database schema has been created using the English terms associated with the ontology concepts. It was, practically speaking, not possible to define a mediator for the original subsystems on the one hand, while it was more practical to have an integrated database (albeit with duplicated data) on the other. As a side effect, the quality of the data has been drastically

[15] Interesting within this perspective is the remark that only 2 out of 22 industrial ontologies surveyed « had clear inference requirements for which knowledge-based systems technology is necessary » [5: p.79].

improved. A lot of manual and semi-automated cleaning had to be performed to standardise the entries of specific table-fields, e.g., the column containing the contact names of CORDIS-FP4 proposals, which had an empty entry, or a name as an entry or a entry of the form “name: *<name>* *Tel: <tel>* *Fax: <fax>*, etc… Some cells contained multiple data elements, e.g. inventors, within a single field that were separated by different delimiters.

One of the results from the EU OntoWeb thematic network, for which VUB STAR Lab was co-responsible in implementing an ontology based semantic web portal and graphical user interface (GUI) [25], was a generic data model to store the ontology and its instances. This data model has been used for the Innovation Router. It allows transparent retrieval of Innovation Router data. E.g., relationships of a concept instance can be retrieved in a single pass. It guarantees a minimum of maintenance at the data and database levels if extensions or modifications of the ontology are needed.

5 Results

In this section, a general overview of the Innovation Router (section 5.1) is presented followed by a discussion of two potential use cases: one on scientific headhunting (section 5.2.1) and one on patent (opportunities) detection (section 5.2.2).

5.1 Overview

We decided to separate the classification of *topics* from their related *activities* such as persons involved, registered patents, articles published, and project proposals being made. The topics are the research areas defined in the project proposals, keywords extracted from the scientific publications, and phrases extracted from the project proposals, patent applications and scientific publications [16].

When one enters the site (see Fig. 2), the focus is on the *Root* topic. At this level, one can see all structuring mechanisms in parallel at the *Topics*-section. The *Activities*-section indicates all activities related to the current topic. Next to every activity, a number indicates the amount of all registered activities. Browsing through the structuring mechanisms is possible by clicking a *SubTopic* of the current topic, until no deeper levels of structuring are encountered. The *Activities*-section changes dynamically according to the selected topic. By clicking on an activity, a browseable result page is generated, containing all elements of the selected activity, similar to a result page generated by a web search engine. Each element is provided with a short description. Clicking on the elements header will display a full description on an individual page. This full detailed page contains besides the description of the element, all relations of this element with elements of the same or other activities. E.g., an article could be written by several authors who are included in the *Person* activity (inter-activity relationship).

[16] The extraction activities have been performed by Bit Media e-Learning Solution (Austria).

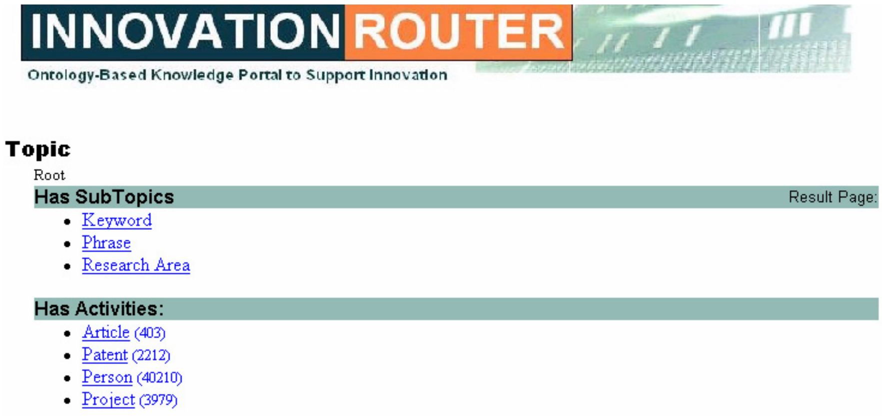

Fig. 2. Innovation Router starting page.

5.2 Use cases

The innovation ontology tells us what kind of things are available in this restricted domain of innovation, how they can be interrelated and what they mean. So first there is the informational need: because the ontology is a structured conceptual model of the innovation vertical domain, it supports parametric search and navigation using product and service knowledge by prospective innovators to discover what to use and to determine their chances and shortcomings. In addition, the ontology maps to the quickly changing data of the competitors. It models not only the product and service knowledge but also knowledge about the end users. By using user personalisation, queries could be customised to the user's experience and status – see e.g. [2]. In the following sections, we present two potential cases of how innovation preparing activities could happen within an enterprise.

5.2.1 Scientific headhunting

One of the competitions between enterprises is the competition for talents. More and more enterprises are aware that high-level talents should be introduced through professional channels, which target the demand of enterprises and find the talents that they need in a timely manner. Recruitment agencies are adopting competency standards, used to classify the acquired competencies of the jobseekers and the required competencies of vacant jobs. Using the data and knowledge within the Innovation Router, we complement the traditional competency-based database search with a thematic search. The thematic search identifies individuals and provides details of and references to their involvement in the aforementioned projects, patent applications and articles, indicating the individual's professional expertise area(s) – see Fig. 3. Currently, considering the nature of the available data within the Innovation Router, thematic search results will be limited to scientific headhunting purposes only.

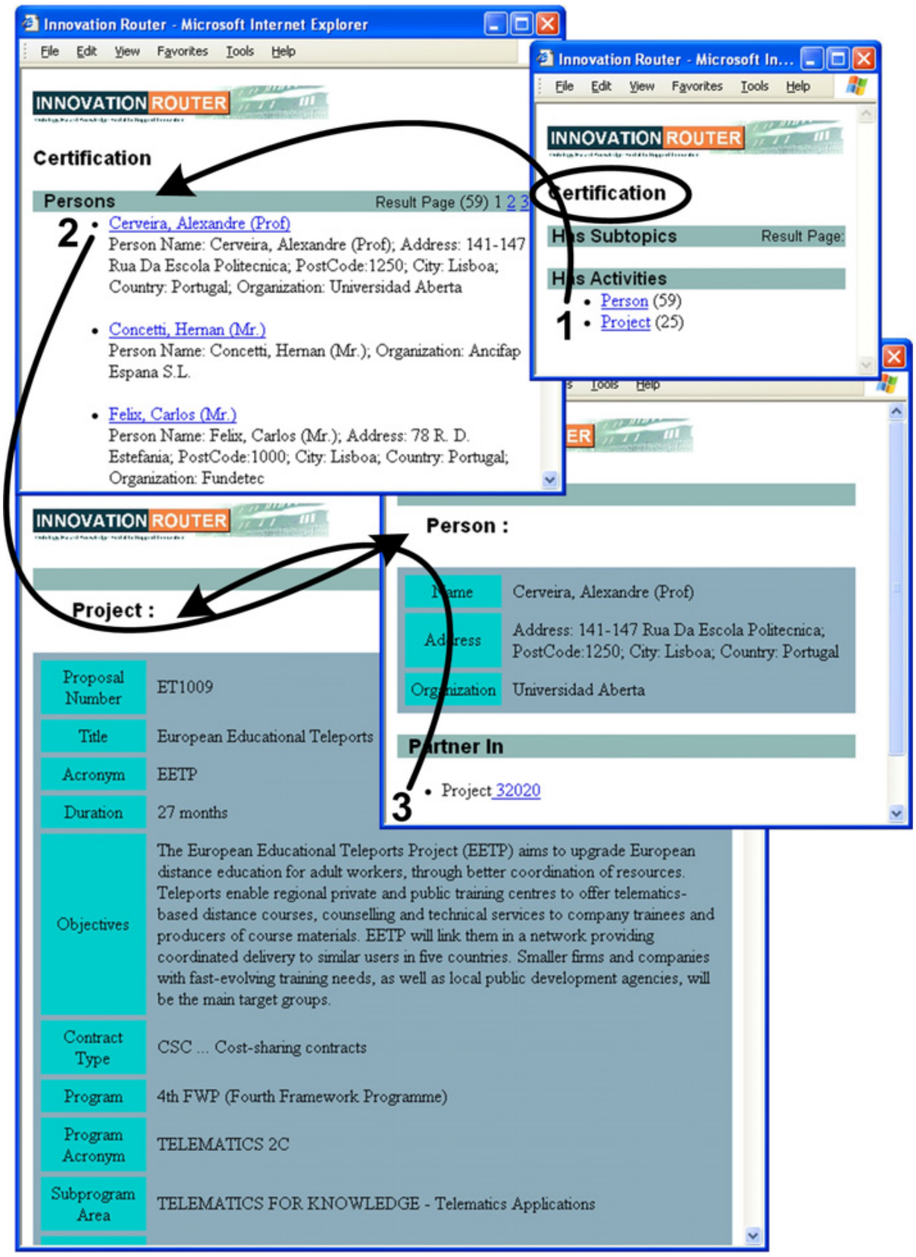

Fig. 3. Performing a thematic search on "certification" returns 59 persons of which one is selected for further reviewing of his contact information as well his professional involvement in various projects, patent applications and articles.

5.2.2. Patent detection

A patent is a right, granted by the government, that excludes others from making, using, or selling the invention covered in the claims of the patent [17]. A patent offers a legal ground for stimulating innovation, where the patentee acquires the right to forbid others to exploit his or her protected invention without permission. Nowadays, patents have become the primary intellectual property asset that companies rely upon to protect their innovations and to maintain competitive advantage by hampering the activities of current and future competitors [22].

Enterprises and individuals require efficient and effective tools and methodologies to identify existing patents and to screen patent opportunities [23]. Within the Innovation Router, we developed a service that contributes to the process of investigating research or business areas in order to identify patent opportunities. Given two themes, besides displaying activities proper to each theme individually, this service calculates the activities overlap between both themes – see Fig. 4. Focusing on the patent-overlap, we are able to present all patent activities which are related to or dependent on both themes. Two themes producing low patent-overlap are potential candidates of technologies, business/research areas, features… to be combined for investigation of novel patent opportunities.

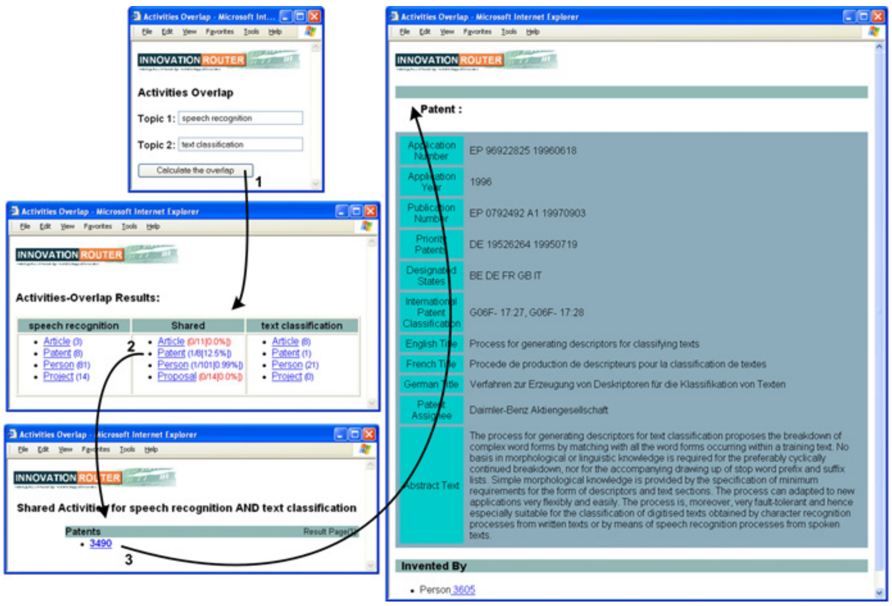

Fig. 4. Given two distinct themes, "*speech recognition*" and "*text classification*", the service displays the activities of both themes individually and calculates the activities overlap between both themes. If overlapping activities do occur, the service allows individual selection of these activities in order to perform a detailed analysis of each item.

[17] http://www.european-patent-office.org

6 Discussion & Related Work

6.1 The Router

The Innovation Router proved the strategic role it could play as part of an innovation process in linking the materials collected for surveys and state-of-the-art studies, in sharing knowledge and emerging insights after confronting the database with enriched queries. While the data available within the framework of the Innovanet project was limited in quantity, it is representative enough to define a preliminary model of innovation or scientific discovery – e.g., to find out who's who and who is doing what – to describe an innovation engineering environment and validate it, to predict trends and to propose potential areas of innovation to an R&D department. An optimised Router enables viewing the world of topics that have activities and systemising innovation services that help decision makers who direct innovation. Furthermore, the Router can be extended into a real strategic tool or instrument in any future research for the systematisation of innovation.

The Innovation Router is clearly related to work on semantic portals (e.g., [4], [18], [25]), schema integration, heterogeneous and federated databases (e.g., [1], [37]). We are currently unaware of similar initiatives targeting in particular the innovation domain, combining insights of the above mentioned research areas and building on accumulated experience in these fields.

6.2 The Modelling Methodology

Existing research methodologies and industry practices cover specific aspects (see [7] for an overview on ontology development; [11] for ontology consistency checking). Currently there hardly exist, at least to our knowledge, comprehensive cookbooks or methodologies (based on one formal and scientific framework) that covers how to actually create from scratch and deploy a *multilingual* ontology-based application. One example are the ONIONS and ONIONS-II [8] methodologies that have been successfully applied to several domains (bio-medical, legal, fishery). They are independent from a particular formal language, but both assume first-order logic, some classification service, and a foundational ontology as a unifying paradigm. Although they are not yet completely engineered, they contain nearly complete guidelines on how to start from scratch and/or to reuse existing sources. Many existing ontology engineering methods build on the CommonKADS [26] knowledge engineering methodology (e.g., [33]) and/or are based on questionnaires for typical expert knowledge elicitation [13]. Others, e.g., Methontology [6], try to encompass the entire knowledge life cycle, but do not provide detailed but generic guidelines (cook book style) on how to construct a domain ontology, as is the case with the adapted ORM/NIAM method.

7 Future Work

More efforts have to be spent on the GUI and human computer interface (HCI) aspects of the Router: e.g., a better visualisation of the search results, more flexible ways of navigation. Some (rather trivial) tweaking of the data needs to be done as well (e.g. formatting). A first step would be to migrate the data and ontology to the OntoWeb semantic portal to benefit from a more flexible and richer GUI. Additional user tests should be performed in order to specify innovation services based on innovation detection patterns. These could be implemented as (semantic) web services. E.g., it will be interesting to discover overlapping activities or expertise between research areas and implement this specific search as a built-in service available to software agents. On the modelling part, the methodology must be applied to other domains, in other circumstances, for other aims, and with a specific eye on the collaborative aspects.

8 Conclusion

The Innovation Router can be perceived as a fundamental element in any strategic innovative thinking. Its inherent innovative characteristic consists of further developing and exploiting recent work on ontologies, thereby offering a multi-dimensional perspective on the relationships between elements of the innovation processes. The parallel representation of data is a powerful means for decision-makers to identify advantages, follow-up actions, gaps and needs. Within the development of the strategic roadmap for the Innovanet project, this Innovation Router was a model or a proof-of-concept. In future applications, this Router can be a stepping stone for the implementation of semantic innovation-related web services, seen its facility to accept enriched queries.

Acknowledgment
This research has mainly been financed by the Innovanet project (EU FP5 IST 2001-38422). Some parts of this research have been funded by the Flemish IWT GBOU 2001 #010069 "OntoBasis" and the Belgian FWO G.0229.03 "BonaTema" projects. The Technology Transfer Centre Leoben, the Austrian project PROVISO, and Johan Hagman of the EU DG INFSO E2 unit are all gratefully acknowledged for providing access to patent data, FP4 and FP5 project information. We also like to thank our colleagues of the consortium for their fruitful collaboration in this project.

References

1 Bergamaschi S., Castano S., Vincini M. & Beneventano D., (2001), Semantic integration of heterogeneous information sources, Data & Knowlegde Engineering 36 (3): 215 – 249.
2 De Bo J., Jarrar M., Majer B. & Meersman R., (2002), Ontology-based author profiling of documents. In, Catizone, R., (ed.), Workshop Proceedings Event Modeling for Multilingual Document Linking (LREC2002 Workshop), pp. 23 – 38, Paris.

3 De Bo J. & Spyns P., (2003), Creating a "DOGMAtic" multilingual ontology infrastructure to support a semantic portal. In R. Meersman, Z. Tari et al., (eds.), *On the Move to Meaningful Internet Systems 2003: OTM 2003 Workshops*, LNCS 2889, page 253 – 266, Springer.
4 Dhraief H., Nejdl W., Wolf B. & Wolpers W., (2001), Open Learning Repositories and Metadata Modeling, In, Proc. of the 1st International Semantic Web Working Symposium (SWWS01), pp. 495 – 514.
5 Ellman J., (2004), Corporate Ontologies as Information Interfaces, IEEE Intelligent Systems jan-feb 2004: 79.
6 Fernández M., Gómez-Pérez A., & Juristo N., (1996), Methontology: From Ontological Art Towards Ontological Engineering, In Workshop on Ontological Engineering (ECAI 96), pp. 41 – 51.
7 Fernández-López M., Gómez-Pérez A., Euzenat J., Gangemi A., Kalfoglou Y., Pisanelli D., Schorlemmer M., Steve G., Stojanovic L., Stumme G., & Sure Y., (2002), A survey on methodologies for developing, maintaining, integrating, evaluating and reengineering ontologies, OntoWeb deliverable #D1.4, Madrid.
8 Gangemi A., Guarino N., Masolo C., Oltramari A., & Schneider L., (2003), Sweetening Ontologies with DOLCE. Proceedings of EKAW 2002. Siguenza, Spain.
9 Gruber T., (1995), Towards Principles for the Design of Ontologies Used for Knowledge Sharing, International Journal of Human-Computer studies, 43 (5/6): 907 – 928.
10 Guarino N. & Giaretta P., (1995), Ontologies and Knowledge Bases: Towards a Terminological Clarification, In, Mars N. (ed.), Towards Very Large Knowledge Bases: Knowledge Building and Knowledge Sharing, IOS Press, Amsterdam, pp 25 – 32.
11 Guarino N., (1998), Formal Ontologies and Information Systems, In, Guarino N. (ed.), Proceedings of FOIS98, IOS Press, pp. 3 – 15..
12 Halpin T., (2001), Information Modeling and Relational Databases: from conceptual analysis to logical design, Morgan-Kaufmann, San Francisco.
13 Hameed A., Sleeman D. & Preece A., (2002), OntoManager: A workbench Environment to facilitate Ontology Management and Interoperability, In, Sure Y. & Angele J. (eds.), (2002), Proceedings of the OntoWeb-SIG3 Workshop at the 13th International Conference on Knowledge Engineering and Knowledge Management EKAW 2002.
14 Humphreys B. & Lindberg D., (1999), The Unified Medical Language System Project: a distributed experiment, In, Lun K.C. (ed.), Improving access to biomedical information, Proceedings of MEDINFO92, Elsevier Science Publishers, pp. 1496 – 1500.
15 Jarrar M, Demey J., & Meersman R., (2003), On Reusing Conceptual Data Modeling for Ontology Engineering, Journal on Data Semantics, 1 (1): 185 – 207 [LNCS 2800, Springer].
16 Jarrar M., Verlinden R. & Meersman R., (2003), Ontology-based Consumer Complaint Management. In, R. Meersman, Z. Tari et al., (eds.), On the Move to Meaningful Internet Systems 2003: OTM 2003 Workshops, LNCS 2889, Springer Verlag, pp. 594 – 606.
17 Lenat, D, (1995), CYC: A large-scale investment in knowledge infrastructure. Communications of the ACM, 38(11), 33 – 48.
18 Maedche A., Staab S., Studer R., Sure Y. & Volz R., (2002), SEAL - tying up information integration and web site management by ontologies, IEEE Data Engineering Bulletin, 25 (1): 10 – 17.
19 Meersman R., (2001), Ontologies and Databases: More than a Fleeting Resemblance, In, d'Atri A. and Missikoff M. (eds), OES/SEO 2001 Rome Workshop, Luiss Publications.
20 Meersman R., (2002), Semantic Web and Ontologies: Playtime or Business at the Last Frontier in Computing ?, In, NSF-EU Workshop on Database and Information Systems Research for Semantic Web and Enterprises, pp.61 – 67.
21 Miller G., (1995), WordNet: a lexical database for English, Communications of the ACM 38 (11): 39 – 41.

22 Mogee M.E., (1997), Patents and technology intelligence. In, Ashton W. & Klavans R., (eds.), Technical Intelligence for Business: Keeping Abreast of Science & Technology, Battelle Press, Washington, D.C.
23 Mogee M.E., (1997), Patent analysis methods in support of licensing. Technology Transfer Society Proceedings of the 1997 Annual Meeting.
24 Nirenburg S. & Raskin V., (2001), Ontological Semantics, Formal Ontology and Ambiguity, In, Proc. of the 2nd International Conference on Formal Ontology in Information Systems, ACM Press, pp. 151 – 162.
25 Oberle D. & Spyns P., (2003), OntoWeb – Knowledge Portal, In, Staab S. & Studer R., (eds.), Handbook on Ontologies, International Handbooks on Information Systems, Springer Verlag, pp. 521 – 540.
26 Schreiber A., Akkermans J., Anjewierden A., De Hoog R., Shadbolt N., Van De Velde W., & Wielinga B., (2000), Knowledge Engineering and Management: The CommonKADS Methodology, MIT Press.
27 Schumpeter J., (1942), Capitalism, socialism and democracy, New York, Harper and Row.
28 Spyns P., Oberle D., Volz R., Zheng J., Jarrar M., Sure Y., Studer R., Meersman R., (2002), OntoWeb - a Semantic Web Community Portal, In, Karagiannis D., & Reimer U. (eds.), Proceedings of the Fourth International Conference on Practical Aspects of Knowledge Management (PAKM02), LNAI 2569, Springer Verlag, pp. 189 – 200.
29 Spyns P., Meersman R. & Jarrar M., (2002), Data modelling versus Ontology engineering, In, Sheth A. & Meersman R. (eds.), SIGMOD Record Special Issue 31 (4): 12 – 17.
30 Spyns P., (2004), Methods to be used in ontology engineering. OntoBasis Deliverable #D1.6, STAR Lab, Brussel.
31 Staab S. & Studer R., (2003), Handbook on Ontologies, International Handbook on Information Systems, Springer Verlag.
32 Stojanovic L., Stojanovic N. & Volz R., (2002), Migrating data-intensive Web Sites into the Semantic Web, in Proc. of the ACM Symposium on Applied Computing (SAC 2002), ACM Press, New York, pp.1100 – 1107
33 Sure Y., (2003), Methodology, Tools & Case Studies for Ontology based Knowledge Management, PhD thesis, AIFB Karlsruhe.
34 Twiss B., (1992), Managing Technological Innovation, London, Pitman Publishers.
35 Verheyen G. & van Bekkum P., (1982), NIAM, aN Information Analysis Method". In, Olle T., Sol H. & Verrijn-Stuart A., (eds), IFIP Conference on Comparative Review of Information Systems Methodologies, Noord-Holland.
36 Vossen P. (ed.), (1998), EuroWordNet: A Multilingual Database with Lexical Semantic Networks, Kluwer Academic Publishers, Dordrecht.
37 Wiederhold G., (1994), Interoperation, Mediation, and Ontologies, In, Proceedings International Symposium on Fifth Generation Computer Systems (FGCS94), Workshop on Heterogeneous Cooperative Knowledge Bases, vol. W3, ICOT, Tokyo, pp. 33 – 48.