Knowledge exploitation from the web

David Riaño[1], Antonio Moreno[1], David Isern[1], Jaime Bocio[1], David Sánchez[1], and Laureano Jiménez[2]

[1] Departament d'Enginyeria Informàtica i Matemàtiques
Universitat Rovira i Virgili
43007 Tarragona, Catalonia (Spain)
{drianyo,amoreno,disern,jbocio,dsanchez}@etse.urv.es

2 Department of Chemical Engineering and Metallurgy
University of Barcelona
08028 Barcelona, Catalonia (Spain)
laure.jimenez@ub.edu

Abstract. In the framework of Knowledge Management, the Internet can be a valuable source of information to produce new Knowledge. Here, an ontology-based web search system to ease the enterprise managers in the process of discovering new knowledge from the documents in the Internet is introduced. By means of a graphical user interface, the user of the system supplies an ontology in RDF to describe the domain of interest, and sets up some predefined parameters in order to constrain the search corpus. A distributed intelligent process works to achieve the levels of quality and quantity about the results that the user established. Several ideas about how to use this system and its application to seven real domains are also supplied.

1 Introduction

Knowledge Management (KM) can be defined to encompass the strategy, the processes, and the technology employed to enable an enterprise to acquire, create, organise, share, and make actionable knowledge needed to achieve the vision of the enterprise. These complex tasks have been traditionally tackled with the use of *first generation knowledge management* (FGKM) tools that are oriented to the administration of information structures in order to deal with the knowledge that benefits the enterprise. These tools are mainly based on information technology and strategic management. Nowadays, the approach to KM is by means of *second generation knowledge management* (SGKM) tools which are oriented to the development of knowledge structures that are captured from the enterprise experiences. Contrarily to FGKM, SGKM systems do not consider valuable knowledge as something that already exists, but something that emerges from the enterprise activities and that must be managed along a *knowledge life cycle* process that is divided into three sequential steps[1] [8]: knowledge production, knowledge validation, and knowledge integration.

Knowledge production involves the creation of new knowledge, such as new ideas, insights and innovation spawned by interaction between people or groups and the acquisition of knowledge from outside sources. *Knowledge validation* is the next step, during which potential new knowledge is subjected to expert review and processes that test its reliability and value in practice. Knowledge that passes these tests is then integrated, or implemented, within the organisation in the *knowledge integration* step.

One of the main problems in knowledge production is that knowledge is not always easily expressed by means of a formal representation language. This fact makes a distinction between what is called *explicit knowledge* (it can be expressed in words and numbers and shared in the form of data, scientific expressions, specifications, manuals, etc.) and *tacit knowledge* (it is far to be reduced to explicit expressions and cannot be articulated, but it can be learned, acquired or managed).

1.1 Using ontologies for Knowledge Management

Knowledge engineering in KM is closely related to the concept of ontology. In this framework, an *ontology* is a conceptualisation of a knowledge domain in a model that defines a structure of concepts and concept relationships which are relevant to the domain. Ontologies are explicit knowledge structures that can be shared [1].

These structures can be used to represent declarative or procedural knowledge about the domain of interest of an enterprise. *Declarative knowledge* (or know-what) consists of the factual assessments an organisation makes about itself and its capabilities in the context of its marketplace and operating environment. *Procedural knowledge* (or know-how) takes the form of business processes that the organisation executes in the areas of marketing, manufacturing, personnel, R&D, etc.

1.2 An ontology-based Internet search tool

KM is not only related with the internal knowledge of a enterprise, but also with its context in the local, the national, and the world wide market. Declarative and procedural knowledge about competitors, their products, their policies, their success, etc. can be very valuable for the manager of a company. Political changes, the publications of new laws, norms, or standards, the appearance of new materials or the development of new production technologies describe also a sort of knowledge that has to be managed for the benefit of the own company. Nowadays, the Internet is the best window to show and find public information all around the world. Unfortunately, although computer browsers that access this information are efficient in the tasks of web page

---

[1] Other authors define different knowledge life cycles; e.g. Davenport and Prusak [10] consider the steps of knowledge generation, codification and transfer.

search and retrieval, they are not so good in the tasks of processing such information (e.g. page selection, knowledge production, validation and integration). This is so because these search tools use to work with a limited keyword-based description of the domain of interest of the user, and not with a semantic representation of the relationships between these keywords (i.e. concepts) or a description of the concepts in terms of their properties and synonyms.

Here, an ontology-based search system is introduced that explores the Internet to find relevant pages related to the different concepts in an ontology. The retrieved pages are textual instances of the concepts, but conditioned to the meaning of the concept in the whole ontology. The content of the pages related to a particular concept are analysed in order to rank them according to a relevance function which takes into account the properties describing the user desired profile of such concept.

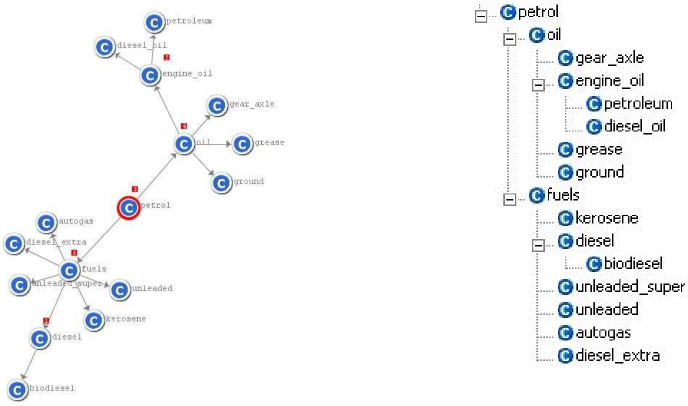
This system is introduced as a decision support tool that can ease and fasten the process of knowledge production and validation by means of a friendly user interface. The system was tested with seven public ontologies: *employment*, *petrol oil*, *biotechnology*, *sustainability*, *axens processes*, *wastewater*, and *technology*. For each one, the system was used for knowledge mining from the web. Particularly, it was used to acquire new explicit knowledge, to validate this knowledge on the basis of the relevance of the pages producing it, and to discover tacit knowledge and make it explicit.

Section 2 describes the ontology model that the search system is able to work with. Then, the system is introduced in section 3 where some ideas about how to exploit the system for knowledge production are highlighted. Section 4 describes the use of the system to produce knowledge in seven alternative application domains. The paper finishes with a section containing some conclusions.

2 Domain representation as an ontology

Knowledge production is the process by which new knowledge is created from a variety of sources: people and groups, reports, intranet and Internet documents, information systems, etc. When the source of information used in knowledge production is textual or oral, natural language processing scientists call it a *corpus*. Knowledge production can be guided by a description of the area of interest defined by the enterprise manager that shows what the manager is willing to search about in the selected corpus. An *ontology* is an explicit formal specification of how to represent objects, concepts and other entities that exist in such area of interest, and the relationships that hold among them. These use to be *class-subclass* relationships. For example, the *petrol oil* ontology in figure 1 describes an area of interest with concepts as *fuel*, *oil*, *diesel* (*fuel*), or *unleaded* (*fuel*) that are related in a taxonomy of petrol products. This could be used by the manager of a petrol company to define the area of interest that filters the knowledge that he wants to obtain from the web pages of a rival company.

Each concept included in the ontology can be refined with the use of *properties*. The properties of a concept describe the sort of features that the manager is looking for. So, for example, *price* is a property of the concept *fuel*, meaning that the manager is interested in the production of knowledge about the cost of fuel. If this ontology was used in a corpus of *stock exchange markets*, the knowledge would be about petrol company stocks, but if the corpus was in *production processes*, the knowledge would be about fuel refinement costs. Concepts and properties can have *synonyms* related.



(*a*) (*b*)

Fig. 1. *Petrol Oil* ontology: (*a*) graphical representation and (*b*) taxonomic representation.

There are many formal languages for representing ontologies [3]: XOL [7], SHOE [5], RDF [4], OIL [2], OWL [4], and others. Here, the RDF notation has been selected because it is simple and flexible, and it allows property inheritance. Some other languages as OIL or OWL sacrifice the RDF simplicity by adding some particularities to deal with more general ontology models. In RDF, concepts are represented as classes, concept properties as attributes, and concept relationships as *subclassof* attributes.

3 An ontology-based search system

Ontologies and corpuses are the main inputs of the web search system. Ontologies represent an area of interest that has been defined by a company manager or selected from an ontology repository (e.g. www.daml.org, protege.stanford.edu, and prise-serv.cpe.surrey.ac.uk). The corpuses are the web search space from which the manager wants to produce knowledge about the domain described by the selected ontology. Once an ontology and a corpus are decided, the search system looks in the corpus for web pages related to the ontology, retrieves them, evaluates their contents, ranks them according to the relevance to the manager interest, and displays them in a graphical interface that helps the user to produce new knowledge.

The system has several control parameters to restrict some aspects of the search as the system response time and response extension, the minimum relevance of the results, etc. All these aspects are described in more detail in the next subsections.

3.1 The search parameters

Before the system starts working, a set of parameters must be supplied. These parameters are divided into two groups: *control parameters* and *corpus parameters*. There are five control parameters (*deadline*, *search engine*, *number of links*, *search agents*, and *threshold*) that can be used to restrict some properties of the search as the expected response time, the search engine used, the response extension, the degree of parallelism of the search process, and the expected quality of the results.

There are also three corpus parameters (*web site*, *language*, and *country*) to define the search space by means of reducing it to all the pages in a particular web site or domain (e.g. *.org*, *.edu*, *.com*, etc.), or restricted to a particular language (e.g. *English*, *Spanish*, *French*, etc.), country or region (e.g. *European Union*, *USA*, *Japan*, etc.).

Table 1. User parameters to customise the web search process.

Parameter name	Description	Default
Deadline	Time allowed to retrieve, rank and sort web pages	20 sec.
Search Engine	Google or Altavista	Google
Number of links	Links to be retrieved for each ontology concept	20
Search Agents	Number of IAs to be used in the search process	1
Threshold	Minimum rate to be attained by a page to be shown in results	50%
Web Site	The search is restricted to the given web site	-
Language	Language in which retrieved web pages must be written	-
Country	Country of origin of retrieved web pages	-

All the parameters have a default value as table 1 shows. In the corpus parameters, the default value indicates that the search is on all the web without any restriction.

3.2 The search engine

The search system uses the APIs of Google and Altavista as search engines wrapped by an intelligent system that cuts the domain ontology into pieces (sub-ontologies) according to the degree of concurrency indicated by the *search agents* parameter, and scatters these pieces between the available search processes running in the computers involved in the process. Then, each search process works to obtain as many relevant pages as the parameters *number of links* and *threshold* respectively indicate. The search engine API used is the one indicated by the *search engine* parameter.

There is a component of the system that predicts whether the number of pages for a particular sub-ontology is adequate. This component works in parallel to the rest of the system in order to modify sub-ontologies with an inadequate number of pages. The process generalizes these sub-ontologies in such a way that the number of pages retrieved can be complemented with some other new relevant ones.

The contents of the pages retrieved are analysed and the relevance of the page is calculated in terms of the sub-ontology concepts and properties appearing in the documents. The relevance is used to rank the pages and to discard those pages with an unacceptable quality. Any text file format is analysed (html, doc, pdf, ps, txt, etc.).

The overall process is restricted by the *deadline* parameter and it stops when this time is reached. This means that the most restrictive the time is, the higher probability there is to have less pages than required.

3.3 The relevance of the results

A concept in an ontology is defined by a description and a context. The description of a concept is represented by the concept *name* (and synonyms, if there are) and the concept *properties* (and synonyms). The context of a concept is represented by the *names of the ancestors* of the concept (and synonyms) and the *inherited properties* (and synonyms) of the ancestors. For example, the description of the concept *kerosene* is {kerosene, kerosine} for name and synonyms, and {flash point, duration} for properties, and the context is {*fuels*, *oil*, *petrol*} for the ancestor names and synonyms and {octane index} for the inherited properties.

Whereas the web pages retrieved by the search engine for a concept only depend on the names of the concept and the ancestors, the quality of these pages is computed with all the properties, inherited or not, of the concept.

Equation 1 shows the quality function that measures the relevance of a page, where p stands for the evaluated page, and A_c for the set of properties of the concept c.

$$R(p, A_c) = \frac{number\ of\ \ properties\ in(p, A_c)}{total\ number\ of\ \ properties(A_c)} \times 100 \tag{1}$$

This evaluation takes into account only the presence or the absence of each property, but it does not count the number of times that the property appears.

When we tested the system we noticed that during the analysis of concepts with a certain depth in the ontology (4-5-6 levels) it was quite difficult to find web pages containing a high percentage of the properties. For this reason, the above function is normalised with respect to the greatest quality obtained for a concept of the ontology, as equation 2 indicates.

$$\overline{R}(p, A_c) = \frac{R(p, A_c)}{\max\limits_{c} R(p, A_c)} \tag{2}$$

3.4 The uses of the search system

Although the search system as it is now is not able to produce new knowledge, it can be used to ease the process of knowledge mining that a company manager may require, and therefore it belongs to what is called *business intelligence*. The system has two potential approaches to knowledge production: one which is based on the analysis of the quantity of links retrieved for each concept and their relevance, and another one which is about the interpretation of the contents of the retrieved pages. As the second approach is still not completely solved in our system, only the possible practical uses of the first approach will be reported here. Two alternative points of view are presented in the next paragraphs. On the one hand, the possible uses of the system when the input (i.e. ontology and parameters) and the output (retrieved pages) of the system change. On the other hand, the possible uses of the system from the perspective of the company department using it.

Ontology: According to the sort of concepts in an ontology, the system is able to help the user in the production of declarative knowledge, procedural knowledge, or both. If the concepts in the ontology represent static concepts as company products the system will help the user to produce declarative knowledge, but if the concepts in the ontology represent dynamic concepts as company processes, then the knowledge produced will be procedural.

Parameters: the wise use of the corpus parameters will permit the user to focus the production of knowledge in the desired context: personal, departmental, in a company, local, regional, national, or international. For example, the system can be used to compare the laws that concern the domain of interest of a company for different countries, and help a manager to decide where to open a new factory.

Retrieved pages: the number of pages retrieved for each concept and their normalised relevance can supply much information about the domain of interest. For example, if no pages about one concept are retrieved from the web pages of a company or the pages retrieved are irrelevant, this reinforces the tacit fact that this company does not produce, sell or offer the product, service or technology that the concept represents. On the contrary, if there are many web pages about one concept of the ontology, the product, service or technology that the concept represents is an important explicit element of the analysed company. When this simple idea is extended to a whole ontology, the implications are very interesting. So, for instance, if a company is in the process of merging another one and the manager has an ontology representing the domain of what the merger is expected to be, then if both companies offer the same products, services or technologies this means that the union of both companies will become a stronger competitor in the same marketplace. On the contrary, if the pages retrieved for each one of the two companies are related to different concepts in the ontology, then the companies complement one another, and the unification will produce the growing of the marketplace. This functionality can be extended to help enterprise managers to find and decide about the company that better fits their needs, according to the interests of the manager represented by an ontology.

Company department: the system can be applied differently according to the interest of the departments of a company. So, in the department of Human Resources it is possible to generate knowledge about where the best professionals to cover the company vacancies can be found (country, university, etc.), or what are the alternative hiring procedures or models. In the Marketing department the manager can produce knowledge about market maps related to the ontology of the company products or services. The study can start with all the internet were the most relevant geographic zones or countries can be detected. Focused on these zones or countries, the most relevant companies, organisations, universities, etc. can be detected, and from them the departments involved in such products or services and even the persons related to them. Finally, for the Production and R&D departments it is also possible to use the system to detect new products or technologies. For example, if we have an ontology with a parent concept representing a sort of product or technology and the sons representing particular products or technologies, we can use the system to search every month about this ontology in the web. If we detect that the number of pages retrieved for the parent concept has increased significantly, and the sum of the retrieved pages for the son concepts remains the same or is reduced, then a feasible explanation that has to be studied in detail is that some new product or technology has appeared replacing the previous ones.

These uses of the search system are a sample of possible applications to help managers to produce knowledge in their domains of interest. In the next section, some experiences about the use of the search system to produce new knowledge in concrete domains are described.

4 Applications

The search system has been tested for producing new knowledge in seven domains of interest: *employment* in chemical engineering, *petrol oil*, *biosensors*, *technologies* in chemical engineering, *processes*, *wastewater*, and *sustainability*. These ontologies have been obtained from public repositories in the internet as the DAML repository at www.daml.org, the Prótegé repository at protege.stanford.edu, and the hTechSight repository at prise-serv.cpe.surrey.ac.uk [6], and some of them have been modified to permit a proper check of the system.

Employment is a model of the concepts involved in the description of educational and professional sectors where jobs are offered in the area of chemical engineering. *Petrol oil* shows a classification of the sort of fuels and oils derived from the petroleum. *Biosensors* is designed to describe the concepts related to companies and products that work on biosensor technology. *Technologies* is concerned about several areas in the chemical engineering technology as process design or thermodynamics. *Processes* contains concepts about the treatment of petroleum residues. *Wastewater* describes concepts about the available legislations in the EU and in the USA about water cleaning, and also about methods to manage water pollutants. Finally, *sustainability* is about technologies that may be applied in any industry to accomplish better ecological measures, and also about tools and methods to be applied in those technologies.

Table 2. Domain ontology features

Ontology name	No. classes	Avg. depth	Max. depth	Attributes?	Synonyms?	D/P
Employment	23	5.0869	7	No	Yes	*D*
Petrol Oil	17	3.7647	5	Yes	Yes	*D*
Biosensors	19	4.3636	6	Yes	Yes	*D*
Technologies	24	4.6666	6	Yes	No	*P*
Processes	15	3.7333	5	No	No	*P*
Waste water	49	4.8867	6	Yes	No	*D/P*
Sustainability	46	4.6739	6	No	No	*D/P*

For each ontology, table 2 supplies extra information as the number of classes of the RDFS document (i.e. *concepts*), the average and maximum depth of the ontology concepts, whether the ontology has attributes or synonyms, and the kind of knowledge this ontology contains: *D* for declarative and *P* for procedural.

In the rest of this section we describe, using several case studies, how the search system can be used to extract different kinds of knowledge and perform diverse types of analysis of documents related to the concepts of a certain domain ontology.

4.1 Case study 1: *Biosensor companies and application areas in Europe*

Biosensors are analytical tools consisting of biologically active material used in close conjunction with a device that will convert a biochemical signal into a quantifiable electrical signal. They have many advantages with respect to other traditional sensors (simplicity, low-cost, fast response time, etc) and many potential applications (agriculture, horticulture, veterinary, pollution, water contamination, clinical diagnosis, biomedicine, etc.). The area of biosensors evolves very rapidly and the management of the new emerging knowledge is something relevant to satisfy the market requirements for small and medium enterprises. Here, the system introduced in section 3 is used to generate knowledge about the more important biosensor companies and areas of application of biosensors in European countries with a high impact of this technology.

For this case study we concentrate on the concepts of the *Biosensors* ontology defined in the IST project *hTechSight* (prise-serv.cpe.surrey.ac.uk/techsight/). First, we use the system to study the number of results provided on the concept *biosensor company* (which obtains more than 22,500 references in Google) restricted to some country domains in Europe (.uk 763, .de 627, .it 145, .es 82, .fr 52, .fi 58, .pt 10, .lv 1). From the analysis of these results we can notice that "*the United Kingdom seems to be the country with more presence of biosensor companies in Europe*". The system also supplies the names of some of "*the most relevant biosensor companies in the UK: Cybersense Biosystems Ltd, BIVDA, The Generics Group, and MCA Services*".

In a further step, we can search pages related to the whole *Biosensor* ontology on each one of the above companies (i.e. www.cysense.com, www.bivda.co.uk, www.generics.co.uk, and www.mcaservices.co.uk). The ontology contains classes about possible areas of application of biosensors: *health-care*, *industry*, *environment* and *veterinary*. By comparing the pages obtained from each company in each of these classes, we may produce different kinds of knowledge: the areas of application of each company can be made explicit from their web pages (e.g. "*The Generics Group is a technology consulting development and investment organisation focused in health-care biosensors and not in other sort of biosensors*"), the areas in which there is a shortage of manufacturing companies is obtained from tacit facts as the absence of relevant pages in certain ontology classes (e.g. "*there are not relevant companies in the UK related to the application of biosensors to veterinary*"), areas in which different companies are competing (e.g. "*the Generics Group and Pinebridge are two companies competing on health-care biosensors*") or areas in which companies are complementary. In these two latter cases, knowledge is implicit in the distribution of web pages among different classes in the ontology.

4.2 Case study 2: *Petrol companies, research groups and institutions*

Petrol is used to produce some of the materials that have a higher impact on the world wide national and international policies and economies. The companies of the sector have to be aware of the global evolution and also about the strategies of their partners and competitors. Here, we use the search system to find knowledge about state-of-the-art the current most relevant companies, research groups and other related institutions in the areas of petroleum and derivatives.

In this case study, the ontology in figure 1 is used to analyse the world wide companies, research groups and other related institutions, restricted to the domains .com, .edu and .org. With the first analysis we detect several "*well known petrol companies as Texaco, Esso, BP, Shell or Total, and*" also some others "*less known companies as Bharat Petroleum*" (www.bharatpetroleum.com) that cover important topics in the ontology as petrol, fuels, kerosene, unleaded, diesel, grease, engine_oil or petroleum. With the second analysis, knowledge about research groups in the petroleum area is produced. For instance, we found references to the web site of The College of Earth and Mineral Science (www.ems.psu.edu) at Penn State University, the Biology department (www.bio.unc.edu) at North Carolina University, and the Department of Earth and Atmospheric Sciences (www.geo.cornell.edu/eas) at Cornell University. Finally, closed related organisations like *Planet Ark* (www.planetark.org), an ecologist organisation covering several topics about fuels, and *Clean Air* (www.cleanairnet.org) an Asian organisation that promotes technologies for the improvement of the air quality around the world were found under some fuel nodes as petrol, unleaded, autogas, diesel, biodiesel and grease. From these results, we can infer knowledge about "*the organisations working on the effects of fuel and derivatives in the environment*" and "*the sort of concern of such organisations with respect to the concepts in the ontology*". For example, this case study can conclude that the "*ecological organisations are more concerned about the petrol derived fuels than about the petrol derived oils*".

4.3 Case study 3: *Sustainability in environmental-aware companies*

Sustainability is defined as a process or state that can be maintained indefinitely, and which integrates an appropriate balance between a viable economy, protection of the environment and social well-being.

Here, the *sustainability* ontology contains technologies, tools and methods to improve the level of sustainability of a company. These are summarised in environmental risk assessment, analysis of the life cycle of a product or process, ecodesign, and process simulation. The search system in section 3 has been used to find the most frequent web domains. Among those which have been found (.com, .edu, .gov, and some others which are proper to countries), the first two ones are the most frequent.

The *wastewater* ontology includes concepts about legislation on water quality and pollutants. The ontology distinguishes between EU and US legislations. In the EU, four countries have been studied: France, Germany, Ireland, and Spain. In the ontology four concepts about wastewater legislation have been analysed in detail: energy, paper, plastic, textile and sodium. The study shows that "*for all the EU countries, the relative importance order of all the concepts about wastewater legislation is always the same: energy, paper, plastic, textile and sodium, being energy the most important*". Moreover, "*the relative order of importance of the EU countries for all the concepts are always the same: Germany, France, Spain, and Ireland*".

5 Conclusions

In the last years the Internet has become the most important source of information. Companies use it for advertising their products and services, universities and research institutions for showing new emerging technologies and knowledge, and governments for forecasting information about legislation. All this kind of information can be relevant to enterprise managers. In this paper, an ontology-based search system to assist them in the process of knowledge production from Internet has been introduced.

In the context of the concept of organisational learning [9] defined as the ability to learn faster than your competitors, some abilities of the search system are: discovery of new marketplaces, market needs, technologies and products; analysis of the competitors and allies; promotion of investments that yield higher rates of innovation or added value; detect the situation of the own company in the world, etc.

The search system has been tested using seven domain ontologies that defined the area of interest of the knowledge to be produced. Among all the results obtained, only those which are interesting to show the capabilities of the system have been introduced and commented as case studies. In these cases the knowledge produced as a results of the use of the system has been made explicit as sentences in italic.

Another purpose of the paper has been to show the flexibility of the system in the way it can be adapted to the requirements of the enterprise manager: enterprise alliances, market analysis, national and international legislation, etc.

As a result of the tests made, the most important lessons learned have been:

- It is very important to have a detailed characterisation of the area of interest in the domain ontology. Thus, a set of domain experts should have agreed on the structure of the domain and the terminology used.
- It is also necessary to think about the different ways in which a class or an attribute may be named, and use synonyms to describe them.
- The presence of relevant attributes in each class is crucial, since they allow the system to distinguish among good and bad pages.
- Final users of the system have to be experts in the domain, and in the different types of knowledge that can be acquired from the system response.

- The system is most useful when used in an iterative fashion (e.g. first making a general search on the concepts of the ontology, and then possibly refining the search on a particular concept on the web pages of a specific company).
- The system provides interesting results, but in most cases it is the final user who has to analyse them, extract some kind of knowledge and proceed with a more specific search.

The work presented is part of a work in progress and the results preliminary. In the future, the system will be extended with some functionalities for knowledge extraction and integration. Knowledge extraction will be used to automatically analyse the contents of the retrieved documents in order to generate new knowledge that will be optionally added to the domain ontology during the knowledge integration process.

Acknowledgements

This work has been partially funded by the EU Project *hTechSight* (IST-2001-33174, [6]). D. Sánchez would like to acknowledge the support of the *Departament d'Universitats, Recerca i Societat de la Informació* of Catalonia.

References

1. M. S. Abdullah, I. Benest, A. Evans, C. Kimble, Knowledge modelling techniques for developing KM systems, 3rd European Conf. on KM, Dublin, Ireland, 2002: 15-25.
2. D. Fensel, I. Horrocks, F. Van Harmelen, S. Decker, M Erdmann and M. Kein. OIL in a nutshell. Lecture Notes in Artificial Intelligence, Vol. 1937, Springer-Verlag, (2000).
3. A. Gómez-Pérez, O. Corcho. Ontology languages for the semantic web, IEEE Intelligent Systems, 54-56, 2002.
4. A. Gómez-Pérez, M. Fernández-López, O. Corcho. Ontological Engineering: with examples from the areas of KM, e-commerce and the semantic web, Springer Verlag, UK, 2003.
5. J. Heflin and J. Hendler. Searching the Web with SHOE. In AAAI-2000 Workshop on AI for Web Search. (2000).
6. h-Techsight Consortium. Deliverable D4.Sample Ontologies for representative trial domains. http://prise-serv.cpe.surrey.ac.uk/techsight/deliverables.asp, 2002.
7. P. D. Karp, V. K. Chaudhri, and J. Thomere: XOL: An XML-Based Ontology Exchange Language, v0.3, (1999). ftp://smi.stanford.edu/pub/bio-ontology/OntologyExchange.doc.
8. M. W. McElroy, The new knowledge management: complexity, learning and sustainable innovation, Elsevier, 2003.
9. P. M. Senge, The fifth discipline: the art & practice of the learning organization, Currency and Doubleday, 1990.
10. T. Davenport, L. Prusak, Working Knowledge: How Organizations Manage What They Know, Harvard Business School Press, 1998.