Advanced Engineering Informatics 27 (2013) 358–365

Contents lists available at SciVerse ScienceDirect

Advanced Engineering Informatics

journal homepage: www.elsevier.com/locate/aei

Analyzing a firm's international portfolio of technological knowledge: A declarative ontology-based OWL approach for patent documents ☆

Maria Bermudez-Edo [a,*], Manuel Noguera [a], Nuria Hurtado-Torres [b], Maria Visitacion Hurtado [a], Jose Luis Garrido [a]

[a] University of Granada, Department of Software Engineering, E.T.S.I.I., c/Saucedo Aranda s/n, 18071 Granada, Spain
[b] University of Granada, Department of Management, Business School, Campus Cartuja s/n, 18071 Granada, Spain

ARTICLE INFO

Article history:
Available online 16 April 2013

Keywords:
Ontology
OWL
Patent analysis
Ontology-driven engineering
Ontology engineering
Firm innovation analysis

ABSTRACT

Patent databases contain large amounts of information about the inventions and metadata of corporate patents (such as the technological domain they belong to, their applicants, and inventors). These databases are available online but since they do not provide explicit information about the relationships between different patent metadata, it is not possible for computers to automatically process such relationships. Several patent ontologies have been proposed so far in order to provide patent knowledge bases with semantics by merging information from different databases and establishing a common vocabulary. However, previous ontology literature has paid limited attention to the representation of specific relationships among metadata and the design of reasoning procedures that would allow some information not explicitly specified in the databases or ontologies to be inferred. This article proposes a methodological approach for the definition of relationships and reasoning tasks for patent analysis by using patent ontologies, and provides a real illustration of its potential in the context of international flows of research knowledge. This declarative method is based on the formal definition of key patent analysis indicators (KPAIs). The case study analysis is relevant because global competition and the importance of multinational firms in the patent process have resulted in firms not only patenting on their domestic markets but also transferring their patents to other markets and developing patents in different countries. In this context, it is important to analyze the connections between the patenting processes and the international knowledge flows of research and development. More specifically, the paper illustrates the applicability of the proposed methodology by classifying patents into the five patterns of internationalization identified by the Organization for Economic Co-operation and Development (OECD).

1. Introduction

In the current context of massive technological change and growing globalization, "innovation is considered strategically important for the ability of firms to build and sustain competitive advantage thus creating value" [1]. The firm's patents are resources that may create a barrier to protect the income generated from innovation efforts or otherwise prevent rivals from competing in that innovative space [2]. In this context, the number of patents filed worldwide grew by 7.2% in 2010 compared to a 5.2% increase in global gross domestic product and over 2 million new patent applications were received in the patent offices worldwide in 2010 [3].

The patent applications are known to contain detailed information such as the applicant's name (or patenting firm), technological type, inventor, geographic location, which is particularly useful for economic development and business research. However, much of the useful information and knowledge that patents can provide requires sophisticated analysis through the connection of multiple data dimensions which are often not explicitly represented in patent databases. This situation generates serious difficulties for the automatic processing of information by computers and traditional patent analysis is therefore expensive and costly in terms of time and manpower [4]. In order to fully exploit patent metadata information and preprocess it for business analysis, this paper proposes an ontology-based method for exploiting the potential of identifying relevant relationships in order to broaden the information and knowledge gathered in patent metadata. The contribution of this paper is to extend previous work on patent ontologies with a declarative approach by means of the formal definition of key

☆ The authors thank anonymous reviewers for their helpful suggestions in this study.
* Corresponding author. Tel.: +34 958243708; fax: +34 958246222.
E-mail address: mbe@ugr.es (M. Bermudez-Edo).

1474-0346/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.aei.2013.02.003

patent analysis indicators (henceforth referred to as KPAIs) in the OWL language.

In recent years, ontologies have become more and more popular in information system engineering [5–9] for describing information as they provide formal, uniform and shareable representations about a domain, and various patent ontologies have been proposed. Patent ontologies classify concepts (such as metadata found in most patent databases or keywords found in patent documents) hierarchically so that patent classifications may be developed and complex relationships between concepts may be examined.

One of the most used ontology languages is OWL [10]. Much of the power of the OWL ontology, whereby model-theoretic semantics are based on description logics, relies on its reasoning capabilities [11]. Reasoning in OWL allows certain verification procedures to be executed such as consistency checks, concept satisfaction and classification. More specifically, when relevant relationships are identified and represented between concepts in OWL, different external reasoners are available to classify concepts based on these relationships.

Our work offers a formal definition of key patent analysis indicators (henceforth referred to as KPAIs) and proposes a methodology to identify and represent new relationships between concepts that allow KPAIs which are not explicitly specified in patent metadata to be inferred. The declarative nature of the proposed methodology involves first declaring the concepts and relationships which are then used by the reasoners to automatically reclassify all present and future data.

By using a sample implementation, our paper also shows how the proposed methodology allows OWL ontologies to represent new relationships between concepts gathered in patent documents. In particular, this research analyses information about international patent flows by implementing a classification which depends on patent internationalization. Such developments also illustrate the potential of ontologies for managing and researching the field of business.

Although previous literature has used artificial intelligence algorithms in the context of ontologies to analyze patents for detecting emerging information from specific keywords [12,13], these ontologies are not using declarative approaches. The declarative approaches complement artificial intelligence algorithms by a priori delimitating concepts in the ontology. Certain ontologies in fields other than patents have identified relevant relationships between concepts to infer new knowledge such as in the field of biological science [14] or in the field of context modeling [15]. As far as we know, none of the previous ontologies has proposed a methodology to identify new relationships and implement them in OWL to infer new knowledge, and no previous work has implemented relationships specifically on the patent data domain.

The information inferred from the proposed methodology could be used by academics and practitioners alike to improve their studies based on patents. This topic is especially relevant in the current context of competition in which firms are increasing their innovative activities organized at international levels [16]. Our case study can lead to the implementation of strategies based on the internationalization of innovative activities, and will allow firms to create a global competitive advantage [17].

The remainder of the paper is organized as follows: Section 2 provides a background study of patents, ontologies and patent ontologies; Section 3 describes the methodology of the proposal and presents a case study; Section 4 complements our methodology with the description of its potential to modificate the declarative approach; Section 5 shows the performance of our proposal by running two reasoners in the context of our case study; and finally, Section 6 concludes the paper by discussing the contributions of the research.

2. Background and literature survey

2.1. Patents

Patents are legal documents that describe technological inventions and protect the rights of the inventor of an industrial property to exploit the invention. Patent documents are organized into multiple metadata fields that contain information such as the inventor, applicant, international technological classification, title, description and claims, among others.

Patents are frequently used as indicators of technology and predictors of economic performance [18]. They are also used to measure the inventiveness of countries, regions, firms or inventors [17]. One of the many uses of patent statistics is to measure the dynamics of the innovation process (e.g. co-operation between industries or countries). These enable the patterns and intensity of international co-invention and foreign ownership of domestic inventions (and vice versa) to be tracked. This is particularly useful nowadays in the global market where it is common to use alliances between different geographical locations to obtain research synergies and to acquire new technological competences [16]. In conclusion, as stated in the OECD's patent statistics manual: "patent data provide unique insights into the processes and outcomes of inventive activities (e.g. the location of inventive activities, inventive networks, emerging technologies, etc.). Used with other data, they support the analysis of other dimensions of innovation that are of policy interest, such as the role of intellectual property in economic performance, entrepreneurship, and the tracking of linkages in the science and technology system" [16].

Traditional patent analyses requires significant cost, time and manpower [4], because they need to download the target patent documents into a customized database and manually query the database in order to reclassify patent documents based on a new classification. Whenever new patent documents must be added to the analysis, the customized database must be updated and a new query must be designed and performed".

Algorithmic approaches are useful to analyze the big amount of information in the patent database. Multivariable analyses (e.g. factorial analysis) are usually used to extract key information without a priory fixed model, and a model is trained with the extracted information to predict consequences [4].

In declarative approaches, the model is a priory proposed, and subsequent changes only need the change in the ontology while software tools may remain stable. In any case, the information and data that patents can provide require sophisticated analysis to connect multiple data dimensions and these are not explicitly represented in any patent database. As a result, it is extremely difficult for computers and statistical software applications to automatically process the information. This paper proposes the development of a methodology used in a patent ontology to declare concepts and new relationships to automatically infer information that is not explicitly written in the patent metadata.

2.2. Ontologies

Various definitions of ontology have been proposed in the literature [19,20] and one of the most widely used is the '*specification of a conceptualization*' [21]. Ontologies consist of a formal description of concepts (classes) and their properties, relationships, constraints and behavior [19,22]. Ontologies describe a common vocabulary for each particular domain, enabling for example the automatic search for concepts, regardless of the terms used to represent the same concepts.

Ontologies are useful when dealing with large collections of information found in data repositories since they are able to maintain the information in a hierarchical organization of concepts and their subsumptions, like a taxonomic hierarchy [23]. In terms of ontology languages, OWL [10] has become the *de facto* standard in the system and software engineering community.

OWL has a formal semantics based on description logics [24]. There are several OWL-based reasoners such as Pellet [25] or Hermit [26] that perform verification procedures such as consistency checks, concept satisfaction and classification.

There are several published proposals for creating ontologies for any purpose by declaring concepts and properties (see [27] for a review) with the most cited being 'Ontology Development 101' [19]. No previous publication, however, deals with the methodology for creating axioms and exploiting the ontology with the power of reasoning in order to enrich the way firms could analyze their innovation.

This paper will use OWL-based reasoners and the most widely used visual tool to create ontologies, Protégé [28] to discuss and illustrate how ontologies may be efficient systems for representing patent information and reasoning, by identifying new contexts and dependency information not shown in the patent databases.

2.3. Patent ontologies

Different ontologies about patents have been implemented in recent years. These ontologies allow different patent databases to be integrated with a homogeneous representation of semantics. Two main different groups of patent ontologies can be identified: patent ontologies that make use of artificial intelligence algorithms for data-mining and patent ontologies that make use of techniques that are based on the declaration of concepts.

In the first group are the ontologies that use the frequency of certain words found in patent documents to analyze technological trends or other useful information [29–31]. This type of work uses text data mining and intelligent algorithms to create patent ontologies. In the ontologies using artificial intelligence, patent data is analyzed and a model is derived from the data.

Other patent ontologies that use artificial intelligence algorithms have been developed with OWL from keywords included in the patents but not from the metadata [32–34]. These ontologies mainly search specific topics to reflect new technological developments and fields but do not use patent metadata such as the applicant, and inventor, which may be useful for business analysis. These ontologies also reclassify patent documents by querying the ontology. However, they do not provide a methodology to implement complex analysis where auxiliary classes and properties are needed. These complex analyses are difficult to implement with queries and it is also difficult to reuse the auxiliary classes in other future complex analyses. In any case, the ontologies based on keywords and other patent information could be combined with patent metadata ontologies to increase the knowledge base, and enabling the information required for business analysis to be processed automatically.

In the second group, ontologies are structured models of concepts to represent the patent metadata (patent ontologies using concept declaration). Although these metadata-based patent ontologies represent the concepts of patent metadata, they do not identify relevant or complex relationships between these concepts in the ontology. The most relevant declarative patent ontologies based on patent metadata and represented in OWL are Patexpert, which is an ontology created within the European Patexpert project [35,36], and the PatentOntology, which was created at Stanford University [37].

Patexpert is a huge patent ontology developed over the core Suggested Upper Merge Ontology (SUMO). The ontology combines several smaller ontologies, such as PMO (patent metadata ontology) that represents all the metadata used in most patent databases. Within the project, several applications using the ontology were developed. The public ontology is not populated and mostly focuses on the delimitation of a global structure to include patent metadata information.

PatentOntology is a newer ontology created explicitly to integrate heterogeneous domains [37] and only for patents recovered from the USPTO (United States Patent and Trademark Office) database. PatentOntology simultaneously represents patent metadata and information from patent court cases. This combination of information helps companies to find not only the patent claims in a technological area but also to discover whether the claims are still valid or have been invalidated by a court. This ontology reduces the time required by companies to search for this information since all the information is in one knowledge base.

The methodology that we present in this paper is based on the reasoners' ability to infer information rather than on the creation of an ontology, as in previous work. More specifically, we use the reasoners to infer new knowledge based on the relationships between ontology concepts. Our proposal could therefore be applied to an ontology created by a declarative approach or on an ontology created through data mining analysis.

3. A proposal enabling explicit relationships to be made between concepts in patent ontologies and a case study

This methodological proposal is based on the declaration of concepts, properties and relationships between concepts, extending the creation of ontologies proposed by Noy and McGuinness [19] and further developed by Darlington and Culley [38] with further steps to declare classes and relationships for inferring new further knowledge with existing reasoning tools. In order to apply this method for analyzing relationships in ontologies, it is desirable to have an already implemented ontology, which we call the '*host ontology*'. In this case, the proposal has been implemented on PatentOntology, although it could be implemented on any existing or new patent metadata ontology.

When the number of international patents filed under the Patent Cooperation Treaty (PCT) grew by 5.7% in 2010 [3], managers had to consider the specific geographical location of their firms and to decide on the specific location of the invention process. It is therefore important to analyze the connections between patenting processes and international knowledge flows of research and development. In this context, the importance of developing specific relationships between concepts to fully understand the situation of this internationalization is particularly important.

In order to illustrate the applicability of the methodology, this paper implements the proposal by classifying corporate patents into the five patterns of internationalization identified by the OECD [16].

The OECD classification aims to study international flows of knowledge based on patent owner country and inventor country, thereby sorting patents according to whether the research has been conducted by subsidiaries of a multinational corporation, by joint ventures or by co-operation between companies, etc. The OECD classification therefore divides patents into five groups:

1. Purely domestic ownership of foreign inventions (country A owner and country B inventor).
2. Domestic ownership implying co-ownership with a single inventor (countries A and B owners but only country B inventor).
3. Domestic ownership with co-invention (countries A and B inventors but only country A owner).
4. Co-ownership jointly with co-invention (countries A and B inventors and owners).

5. Cross-border ownership or inventorship with separate inventor and owner countries (A and B owners and C inventor).

This OECD classification of the geographical location of innovative activities is important because of its implication for the technological capacity of countries and in order to quantify the intensity and geographical patterns of these activities [39,40].

The following subsections describe each of the steps comprising the proposed method (see Tabel 1) and the method is also implemented as an example. The example is the OECD classification in a patent ontology. This example is implemented on PatentOntology (which acts as the ‘‘host ontology’’) and highlights how by following our method reasoners may classify current and future patent documents into the corresponding OECD groups.

3.1. Step1. Enumerate the key patent analysis indicators (KPAIs)

3.1.1. Procedure

Developing an ontology begins with the delimitation of the important concepts in the ontology domain. This involves not only a conceptual mapping whereby experts in the target ontology domain identify the domain’s important concepts through brainstorming or discussion but also the definition of concepts and identification of synonyms or terms defining the same concept. The first step in our proposed methodology is similar and consists of the full delimitation of new concepts delimitated by different concepts of the ontology: these new concepts are the key patent analysis indicators (KPAIs) for our analysis. The host patent ontology includes metadata information such as applicant, inventor and applicant country, and the KPAIs will be defined by specific relationships between these different metadata in the ontology.

3.1.2. Application

This step enumerates the important concepts to deal with our objectives by identifying the KPAIs delimitated by the OECD classification. The KPAIs will be the different patterns that classify patents into an international research flow classification depending on the specific values of the applicants’ and inventors’ countries. In this case, there are five KPAIs, each corresponding to each of the five patterns identified by the OECD:

1. PurelyDomesticForeignInventors: representing purely domestic ownership of foreign inventions.
2. CodomesticSingleInventor: representing domestic ownership implying co-ownership with a single inventor.
3. DomesticWithCoinventors: representing domestic ownership with co-invention.
4. Jointly: representing co-ownership jointly with co-invention.
5. CrossBorder: representing cross-border ownership or inventorship with distinct inventor and owner countries.

3.2. Step 2. Enumerate auxiliary terms

3.2.1. Procedure

This step involves defining the auxiliary concepts which are required to represent the complex relationships (or axioms) that will complete the KPAI definition. These auxiliary terms depend heavily on how each specific analyst sees the structure of the analyzed complex relationships. In their general method for representing ontologies, Noy and McGuinness recognize that: ‘*there is no one correct way to model a domain – there are always viable alternatives*’ [19]. In order to illustrate the different ways of delimitating one specific auxiliary term: If the analyst needs patent delimitation to simultaneously include a specific field and a filing date, it will be possible to use at least two different methods. The first is to delimitate two auxiliary terms (patents in that field and patents with that filing date) and to combine both into a subsequent axiom. The second is to delimitate a single auxiliary term which includes both features. We will illustrate this step with the discussion of our case analysis, where this assertion could also be applied.

3.2.2. Application

In the case study, we delimitate one of the auxiliary terms as the number of countries where the different applicants reside. The OECD classification must use the patent owner for classification purposes. We will use the field ‘‘applicant’’ to represent the patent owner as the applicant is the first owner of the patent and will continue to be so until the patent is sold to a third party. Another auxiliary term which is required for Step 5 is the number of different countries where the inventors reside. These auxiliary terms will help to create the axioms in Step 5.

3.3. Step 3. Represent the KPAIs and auxiliary terms in OWL classes and class hierarchies

3.3.1. Procedure

Once the KPAIs have been identified, this step consists in ordering the KPAIs hierarchically and representing them as ontology classes. There are several ways of ordering the hierarchy: bottom-up, top-down, or a combination of the two. Although all of these could lead to a different class hierarchy, there is no clear better way to represent an ontology and it will really depend on author preferences for each application. Nevertheless, patent analysis through the representation of the KPAIs will normally generate a reduced number of classes to organize and there are no large differences between the methods chosen for ordering the hierarchy. This step also involves the representation of the auxiliary terms defined in Step 2 as ontology classes.

3.3.2. Application

In this step, we represent the concept ‘‘classification OECD’’ with the OWL class `InternationalOECD`, and as a subclass of the parent class `Patent` which belongs to the host ontology PatentOntology. Additionally, the five KPAIs identified by the OECD are subclasses of `InternationalOECD`.

In the same way, the auxiliary terms are also represented as OWL subclasses of the parent class `Patent`, which belongs to the PatentOntology. We represent these classes in OWL as: `NumberApplicantCountries` and `NumberInventorCountries`. Both classes have their subclasses and we shall now represent these in OWL (in brackets): one subclass involving only one country (`ApplicantCountry1` and `InventorCountry1`), another subclass involving two countries (`ApplicantCountry2` and `InventorCountry2`) and the final subclass involving three or more countries (`ApplicantCountry3more` and `InventorCountry3more`).

3.4. Step 4. Define class properties

3.4.1. Procedure

Class properties represent the internal structure of concepts. This step declares the properties that the classes representing each KPAI need. These properties link one class to another or one class with a type of data: integer, dateTime, Boolean, etc. In this step, property restrictions are also described, e.g. domain (classes that a property describes) and range (permitted classes with which a property links).

3.4.2. Application

In order to define the axioms of Step 5 that will allow the KPAIs in this case study to be inferred, certain class properties between patent and applicant countries and inventor countries must be defined. These class properties are:

`hasApplicantCountry`. This property links patents with applicant countries. In accordance with description logic notation [24]

Domain (hasApplicantCountry, PatentDocument)

Range (hasApplicantCountry, State)

`hasInventorCountry`. This property links patents with the inventor countries.

Domain (hasInventorCountry, PatentDocument)

Range (hasInventorCountry, State)

3.5. Step 5. Create the reasoning axioms

3.5.1. Procedure

This is the main step of our proposed method. All the classes and properties represented in the previous steps and the classes and properties existing in the host ontology are now combined so that complex relationships may be made between them. The reasoning axioms will be useful for delimitating relationships to enable patents to be classified according to certain KPAIs.

Unlike traditional databases, OWL follows the open world assumption (OWA) which implies that if something is not explicitly stated then we cannot assume that it is true or false merely that it is unknown. In order to infer new knowledge and reason about certain relationships, it is sometimes necessary to close this OWA [41] by using closures axioms. When these axioms are applied on the properties, they limit the extension of the properties of a class.

This step will combine existing or auxiliary classes or properties to define the classes created in Step 1 (KPAIs). This step also creates the auxiliary closure axioms necessary to implement the reasoning.

3.5.2. Application

Firstly, in order to represent the classes of the OECD classification that involve inventors residing in a single country, two different countries or three or more countries, it is advisable to close the OWA so as to enable reasoning in some of the classes identified in Step 1. This step applies closure axioms over the classes `NumberOfApplicantCountries` and `NumberInventorsCountries` for the six different subclasses (`ApplicantCountry1`, `InventorCountry1`, `ApplicantCountry2` and `InventorCountry2`, `ApplicantCountry3more` and `InventorCountry3more`). For example, the class `ApplicantCountries2` needs a closure axiom so that the cardinality of the property "`hasApplicantCountry`" is exactly two countries. According to description logic notation:

$$InventorCountries2 \equiv \leqslant 2 hasInventorCountry \sqcap \geqslant 2 hasInventorCountry \qquad \text{(Axiom1)}$$

Consequently, all the patents declared as individuals of this class have exactly two applicant countries. These closure axioms will enable us to determine the exact number of applicant or inventor countries and this will subsequently be used to create other axioms. Not only do the closure axioms enable equivalent classes to be represented with negations but they also allow us to represent the exact number of inventor and applicant countries.

Secondly, this step implements the axioms that represent the classes in Step 1. For example, the class `CrossBorder` is a class with two applicant countries and one inventor country where the inventor country is different from those of the applicants. For the sake of clarity in the figure and axioms, the case study has made a simplification here which does not affect the global applicability of the illustration. The case study only includes the top five patent filing countries: USA, Japan, China, Germany and Korea (US, JP, CN, DE and KR). For the complete ontology (with all the countries), it is necessary to increase the number of countries to that defined by the World Patent International Organization (WIPO) in its Standard ST.3 [42].

In description logic, this equivalent class is represented as:

$$\begin{aligned} CrossBorder \equiv\ & ApplicantCountries2 \sqcap InventorCountry1 \\ & \sqcap(((\exists hasApplicantCountry.\{CNi\} \sqcap \exists hasInventorCountry.\{CNi\}) \\ & \sqcup(\exists hasApplicantCountry.\{JPi\} \sqcap \exists hasInventorCountry.\{JPi\}) \\ & \sqcup(\exists hasApplicantCountry.\{KRi\} \sqcap \exists hasInventorCountry.\{KRi\}) \\ & \sqcup(\exists hasApplicantCountry.\{USi\} \sqcap \exists hasInventorCountry.\{USi\}) \\ & \sqcup(\exists hasApplicantCountry.\{DEi\} \sqcap \exists hasInventorCountry.\{DEi\}))) \end{aligned} \qquad \text{(Axiom2)}$$

The implementation of this axiom in Protégé is shown in the upper part of Fig. 1.

3.6. Step 6. Invoke the reasoner

3.6.1. Procedure

This is the final step in which new information will be automatically inferred from the relationships (or axioms) identified above and is performed by OWL language reasoners. OWL (and the updated version OWL-2) enables a domain to be described logically. One of its main advantages is that reasoners for this language can be used to infer further information about this state of affairs [43]. The algorithms beneath these inferences are not part of the OWL document and are mostly implemented by third parties as it is difficult for most organizations to duplicate the assortment of capabilities that reasoners can provide [44]. The reasoners will perform verification procedures such as KPAI consistency checks, concept satisfaction and more specifically classification, which automatically classify instances into the newly created classes in Step 1 or KPAIs.

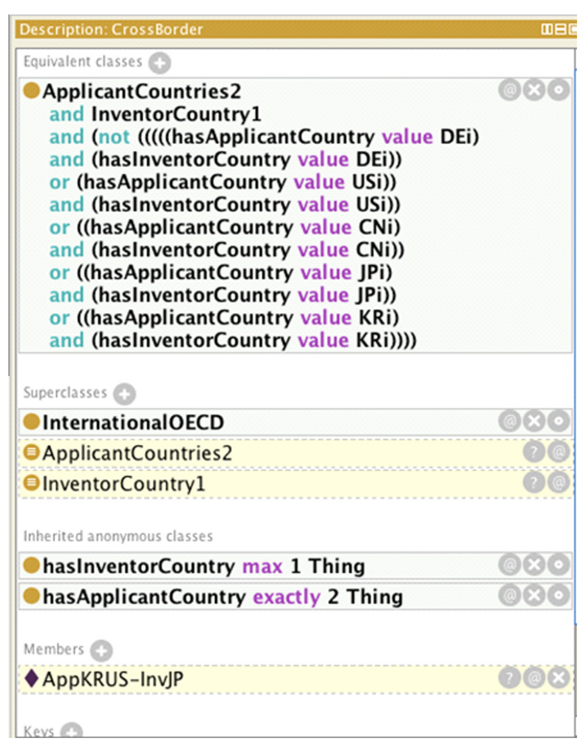

Fig. 1. Class CrossBorder.

3.6.2. Application

In the case study, the reasoner Hermit has classified every patent document into its respective OECD class or KPAI. One example of this can be found in Fig. 1, where the reasoner has inferred that the patent `AppKRUS-InvJP` (a patent with applicants in two countries KR and US and with inventors in one country JP) belongs to the class `CrossBorder`. The remaining patent documents have also been classified with the corresponding KPAIs and the results were similar (specific details are available for the interested reader from the authors upon request). With these declarations of classes, properties and axioms, all the patents that will populate the ontology in the future will acquire the corresponding international classification.

4. Management of changes

The proposed declarative ontology is particularly useful since it allows the classification to be efficiently adapted to future requirements. Our proposal allows an efficient way to implement modifications over the axioms. In a context of changing business and technological interests, the opportunities for a quick adaptation provides a robust alternative to analyze new situations and different interests.

It might be useful to consider different international corporate classifications, such as the one proposed by Rugman and Verbeke [45] or the modifications subsequently proposed by Dunning et al. [46]. Even if the OECD changes its classification, the proposed methodology can be used to change or redefine the axioms, and the current and future patent documents will be reclassified accordingly. In this way, the declarative approach proposed allows patent documents to be analyzed without requiring further programming tasks to be performed. In the current context of increasing patent internationalization, for example, the OECD could be more restrictive in its classification by redefining `CrossBorder` to a class with two applicant countries and two inventor countries (instead of the single inventor country in the old `CrossBorder` definition) and where the inventor countries are different from those of the applicants. In this case, we must redefine `CrossBorder` by simply replacing `InventorCountry1` with `InventorCountry2` in Axiom (2) (see Axiom (3) and Fig. 5).

$$\begin{aligned} NewCrossBorder \\ &\equiv ApplicantCountries2 \sqcap InventorCountries2 \\ &\sqcap (((\exists hasApplicantCountry.\{CNi\} \sqcap \exists hasInventorCountry.\{CNi\}) \\ &\sqcup (\exists hasApplicantCountry.\{JPi\} \sqcap \exists hasInventorCountry.\{JPi\}) \\ &\sqcup (\exists hasApplicantCountry.\{KRi\} \sqcap \exists hasInventorCountry.\{KRi\}) \\ &\sqcup (\exists hasApplicantCountry.\{USi\} \sqcap \exists hasInventorCountry.\{USi\}) \\ &\sqcup (\exists hasApplicantCountry.\{DEi\} \sqcap \exists hasInventorCountry.\{DEi\}))) \end{aligned} \quad \text{(Axiom3)}$$

In this case, instead of classifying the patent `AppKRUS-InvJP` as a `CrossBorder` patent, the reasoner classifies the patent `AppKRUS-InvJPCN` (where Korea and USA are the applicant countries and Japan and China are the inventor countries) as a `CrossBorder` patent.

This proposal highlights the usefulness of representing relevant relationships in OWL in order to increase the knowledge gathered in ontologies or databases. Despite the recognized importance of identifying relevant relationships in the context of patent ontologies, previous publications have used the relationships between concepts in a limited way. The proposal also focuses on ontology specification processes to represent complex relationships in patent metadata ontologies in order to specify certain information and allow the inferred information to be automatically classified and processed by computers. These relationships will provide insights and outcomes of inventive activities [16] and together with other data could show the importance of industrial property strategies in economic performance [47].

5. Evaluation of the reasoning process

In order to evaluate the feasibility and reasonability of the proposed methodology, this section shows the reasoning performance (Step 6 of the proposal methodology), with the two most popular tableaux-based reasoners: Pellet 2.2.0 and Hermit 1.3.5. These two reasoners are used with expressivity OWL-2 DL, and both as plug-ins of the editing ontology tool Protégé version 4.2.0 (build 249) with OWLAPI integrated.

The case study has been built on PatentOntology. Firstly, we have removed all the individuals from PatentOntology and populated it with the controlled individuals that we need for our analysis. Secondly, we have implemented our methodology by adding new classes properties and axioms to PatentOntology.

The complete ontology is written in OWL-2 DL. The main ontology metrics are: it has 79 classes, 45 object properties, 21 individuals (other than patent documents), 73 subclass axioms, 12 equivalent classes axioms (all of them written for this paper), one disjoint class axiom, 65 class assertion axioms and 51 object property assertion axioms. All the experiments are performed on a notebook running OS X 10.6.8 and Java 1.6 on an Intel® Core™2 Duo CPU at 1.86 GHz. Times are taken as the average over 10 independent runs as they were in previous work [48,49].

Table 1 shows the performance of the two reasoners when carrying out the most relevant tasks for our case study. These measures have been taken with different sizes of the ontology ranging from 20 to 50 patent documents. Time is shown in milliseconds (ms).

Table 2shows that performance of the two reasoners is feasible and enables the reasoners to determine all the subsumption relationships without incurring inconsistencies in a few seconds. In comparison with traditional patent analysis, the researcher's time is therefore significantly reduced. In traditional patent analysis, the researcher must manually retrieve patent documents that meet certain requirements (such as OECD classification) and perform the retrieval process whenever new patent documents are published. With our methodology, the researcher only needs to write a few axioms and the reasoners will automatically classify all the patent documents in a reduced time. Furthermore, these axioms will be valid for future patent documents in which case the reasoners will only take a matter of seconds to execute automatic reclassification.

6. Conclusions

Patents are good indicators of the processes and outcomes of inventive activities (e.g. the location of inventive activities, inventive networks or emerging technologies), the technological competitiveness of countries, and the cooperation between institutions, firms and countries in the patenting process. Since patents are stored in large databases where this co-operation is not explicitly connected, it is not therefore possible to automatically process these relationships by computer. Ontology languages such as OWL have proved useful for representing knowledge and creating new relationships which are not explicitly written in traditional databases.

This work highlights the usefulness of identifying and representing complex relationships in ontologies to automatically process information that is otherwise hidden in databases and to declaratively identify key analysis parameters.

The paper first defined the specific steps required to represent complex relationships in OWL by representing the KPAIs in order

Table 1
Methodology proposed in this paper.

Step	Description
1. Enumerate the key patent analysis indicators (KPAIs)	Identify the key patent analysis indicators (KPAIs)
2. Enumerate auxiliary terms	Define the auxiliary concepts
3. Represent the KPAIs and auxiliary terms in OWL classes and class hierarchies	Sort the KPAIs and auxiliary terms hierarchically and represent them as ontology classes
4. Define class properties	Declare the relationships or properties that the classes representing each KPAI need Describe the property restrictions
5. Create the reasoning axioms	Define the KPAIs based on existing or auxiliary classes or properties. This step also creates the auxiliary closure axioms needed to implement the reasoning (OWA)
6. Invoke the reasoner	Automatically infer the relationships (or axioms) identified above. This step is performed by OWL language reasoners

Table 2
Performance of the reasoners in the case study.

Number of patent documents		Pellet		Hermit	
		20	50	20	50
Displayed class inferences	Unsatisfiability	1.3	2.2	0.8	1.3
	Equivalent classes	3.5	7.5	1.7	1.7
	Superclasses	1.7	1.9	1.6	1.8
	Class members	1.8	2.8	1.7	2.2
	Inherited classes	0.2	0.4	0.3	0.4
	Disjoint classes	230.1	677.8	34	304.2

to create new classifications that are not explicitly shown in databases. We also demonstrated and implemented our proposal using an example in which the KPAIs were the international patent classifications proposed by the OECD. In the current competitive context in which firms are increasing their innovative activities on an international level [16], it is particularly important to analyze the internationalization flows of patents. In our case study, the six steps of the proposal are represented. The case study uses the OWL class and property declarations and axioms (including closure axioms) to implement equivalent classes. These classes are then used by a reasoner (Hermit and Pellet in our example) to classify patents into their corresponding KPAI classes which in this particular case are the OECD classification classes.

Our research highlights the flexibility of ontologies when new descriptions of key concepts, or KPAIs, are proposed or needed. Reasoning enables relationships to be efficiently defined in the ontology so that all the information may be automatically reclassified within this new classification in few seconds. A future combination of artificial intelligence algorithms and this reasoning proposal would be especially useful, not only for the analysis of patent metadata, but also to explicitly show the relationships of any data gathered in patents. An effective system to better understand a firm's international patent portfolio is also useful for managers in a context of growing internationalization.

This paper has practical implications for engineers and researchers, enabling a systematic approach for analyzing the complex and extensive patent information and reducing the analysis time by automatizing the classification defined by the axioms that we declare in the ontology. The interest of extending ontologies with the representation of complex relationships is particularly highlighted in this paper. Managers could also benefit from the development of this proposal since it could result in more efficient ways to analyze the firm's patent portfolio and implement the correct patent filing policies. Although this paper presents a methodology for representing complex relationships with patent data, future studies might also add financial data from different databases to the current ontology. With this data combination, managers and researchers could be more effective in identifying how international patenting decisions affect a firm's performance. Although the proposed methodology has been applied to a specific case study, this methodology could be used to any patent analysis, reinforcing its interest and applicability.

Acknowledgments

This research has been funded by the Spanish Ministry of Economy and Competitiveness through R&D Project TIN2012-38600.

References

[1] F. Ciabuschi, H. Dellestrand, O.M. Martín, Internal embeddedness, headquarters involvement, and innovation importance in multinational enterprises, Journal of Management Studies 48 (2011) 1612–1639.
[2] R.M. Grant, The resource-based theory of competitive advantage: Implications for strategy formulation, California Management Review 33 (1991) 114–135.
[3] World Intellectual Property Organization (WIPO), World Intellectual Property Indicators 2011, Available at: http://www.wipo.int/ipstats/en/statistics/patents/.
[4] A.J.C. Trappey, C.V. Trappey, C.-Y. Wu, C.-W. Lin, A patent quality analysis for innovative technology and product development, Advanced Engineering Informatics 26 (2012) 26–34.
[5] A. Weissman, M. Petrov, S.K. Gupta, A computational framework for authoring and searching product design specifications, Journal Advanced Engineering Informatics 25 (2011) 516–534.
[6] H. Panetto, M. Dassisti, A. Tursi, ONTO-PMD: product-driven ontology for product data management interoperability within manufacturing process environment, Advanced Engineering Informatics 26 (2012) 334–348.
[7] A.B. Pelegrina, K. Benghazi, M.V. Hurtado, M. Noguera, A framework for the semantic representation of business processes within business organizational models, Enterprise and Organizational Modeling and Simulation, Lecture notes in Computer Science 63 (2010) 79–94.
[8] G.La. Rocca, Knowledge based engineering: between AI and CAD. Review of a language based technology to support engineering design, Advanced Engineering Informatics 26 (2012) 159–179.
[9] M. Noguera, M.V. Hurtado, M.L. Rodríguez, L. Chung, J.L. Garrido, Ontology-driven analysis of UML-based collaborative processes using OWL-DL and CPN, Science of Computer Programming 75 (2010) 726–760.
[10] P. Hitzler, M. Krotzsch, B. Parsia, P. Patel-Schneider, S. Rudolph, editors, OWL 2 Web Ontology Language: Primer, W3C Recommendation 27 October 2009. Available at http://www.w3.org/TR/owl2-primer/.
[11] M.K. Smith, C. Welty, D.L. McGuinness, OWL Web Ontology Language Guide, W3C Recommendation 10 February 2004, Available at http://www.w3.org/TR/owl- guide/.
[12] A.J.C. Trappey, F.-C. Hsu, C.V. Trappey, C.-I. Lin, Development of a patent document classification and search platform using a back-propagation network, Expert Systems with Applications 31 (2006) 755–765.
[13] M. Shih, D. Liu, Patent classification using ontology-based patent network analysis, in: P.A.C. Systems (Ed.), PACIS 2010 Proceedings. Pacific Asian Conference in, Information systems, 2010, pp. 962–972.
[14] K. Wolstencroft, R. Stevens, V. Haarslev, Applying OWL reasoning to genomic data, in: C.J.O. Baker, H. Cheung (Eds.), Semantic Web Revolutionizing Knowledge Discovery in the Life Sciences, 2007, pp. 225–248.
[15] C. Bettini, O. Brdiczka, K. Henricksen, J. Indulska, D. Nicklas, A. Ranganathan, D. Riboni, A survey of context modelling and reasoning techniques, Pervasive and Mobile Computing 6 (2010) 161–180.
[16] OECD: OECD Patent Statistics Manual. Statistics, vol. 162, 2009.
[17] P. Almeida, A. Phene, Subsidiaries and knowledge creation: the influence of the MNC and host country on innovation, Strategic Management Journal 25 (2004) 847–864.

[18] D. Vries, Leveraging Patents Financially, Gabler Verlag, Wiesbaden, 2012.
[19] N.F. Noy, D.L. McGuinness, Ontology development 101: A guide to creating your first ontology, Technical Report SMI-2001-0880, Stanford Medical Informatics, 2001.
[20] N. Guarino, P. Giaretta, Ontologies and knowledge bases: towards a terminological clarification, in: N. Mars (Ed.), Towards Very Large Knowledge Bases Knowledge Building and Knowledge Sharing, IOS Press, 1995, pp. 25–32.
[21] T.R. Gruber, A translational approach to portable ontology specifications, Knowledge Acquisition 5 (1993) 199–220.
[22] T.R. Gruber, Toward principles for the design of ontologies used for knowledge sharing, International Journal Human Computer Studies 43 (1993) 907–928.
[23] Ghassan Beydouna, A.A. Lopez-Lorcaa, F. García-Sánchezb, R. Martínez-Béjarb, How do we measure and improve the quality of a hierarchical ontology?, Journal of Systems and Software 84 (2011) 2363–2373
[24] F. Baader, The Description Logic Handbook: Theory, Implementation and Applications, Cambridge University Press, 2003.
[25] E. Sirin, B. Parsia, B. Grau, A. Kalyanpur, Y. Katz, Pellet: a practical OWL-DL reasoner, Web Semantics Science Services and Agents on the World Wide Web 5 (2007) 51–53.
[26] R. Shearer, B. Motik, and I. Horrocks. HermiT: A highly efficient OWL reasoner, In *Proceedings of the 5th International Workshop on OWL: Experiences and Directions (OWLED 2008), 26-27 October 2008, Karlsruhe, Germany*.
[27] N. Dhalem, A. Hahn, User-Friendly Ontology Creation Methodologies – A Survey, In Proceedings of the Fifteenth Americas Conference on Information Systems, AMCIS, August 2009, San Fransisco, CA, 1–9. Association for Information Systems Electronic Library (AISeL). Paper 117.
[28] H. Knublauch, R.W. Fergerson, N.F. Noy, M. Musen, The protégé OWL plugin: an open development environment for semantic web applications, LNCS 3298 (2004) 229–243.
[29] A. Trappey, C. Trappey, C.-Y. Wu, Automatic Patent Document Summarization for Collaborative Knowledge Systems and Services, Systems Engineering Society of China, 2009. co-published with Springer-Verlag GmbH.
[30] M.-J. Shih, D.-R. Liu, M.-L. Hsu, Discovering competitive intelligence by mining changes in patent trends, Expert Systems with Applications 37 (2010) 2882–2890.
[31] Y.G. Kim, J.H. Suh, S.C. Park, Visualization of patent analysis for emerging technology, Expert Systems with Applications 34 (2008) 1804–1812.
[32] V.-W. Soo, S.-Y. Lin, S.-Y. Yang, S.-N. Lin, S.-L. Cheng, A cooperative multi-agent platform for invention based on patent document analysis and ontology, Expert Systems with Applications 31 (2006) 766–775.
[33] H. Seol, S. Lee, C. Kim, Identifying new business areas using patent information: a DEA and text mining approach, Expert Systems with Applications 38 (2011) 2933–2941.
[34] S. Choi, D. Kang, J. Lim, K. Kim, A fact-oriented ontological approach to SAO-based function modeling of patents for implementing function-based technology database, Expert Systems with Applications 39 (2012) 9129–9140.
[35] L. Wanner, S. Brügmann, B. Diallo, M. Giereth, Y. Kompatsiaris, E. Pianta, G. Rao, P. Schoester, V. Zervaki, PATExpert: Semantic Processing of Patent Documentation, Knowledge Creation Diffusion Utilization (2009).
[36] M. Giereth, A. St, M. Rotard, T. Ertl. Application of semantic technologies for representing patent metadata, in: 1st International Workshop on Applications of Semantic Technologies AST06, vol. 94, 2006, pp. 297–305.
[37] S. Taduri, G.T. Lau, K.H. Law, H. Yu, J.P. Kesan, An ontology to integrate multiple information domains in the patent system, 2011, in: IEEE International Symposium on Technology and Society ISTAS, 2011, pp. 23–25.
[38] M.J. Darlington, S.J. Culley, Investigating ontology development for engineering design support, Advanced Engineering Informatics 22 (2008) 112–134.
[39] N. Lahiri, Geographic distribution of R&D activity: how does it affect innovation quality?, Academy of Management Journal 53 (2010) 1194–1209
[40] M. Quintas, X. Vazquez, J. Garcia, G. Caballero, Geographical amplitude in the international generation of technology: present situation and business determinants, Research Policy 37 (2008) 1371–1381.
[41] S. Grimm, M. Boris, Closed world reasoning in the semantic web through epistemic operators, OWL: Experiences and Directions (2005).
[42] WIPO: Standard st. 3 recommended standard on two-letter codes for the representation of states, other entities and intergovernmental organizations, <http://www.wipo.int/standards/en/part_03_standards.html>.
[43] P. Hitzler, M. Krötzsch, B. Parsia, P.F. Patel-Schneider, S. Rudolph, OWL 2 Web Ontology Language Primer, second ed., 2012.
[44] M.K. Smith, C. Welty, D.L. McGuinness, OWL Web Ontology Language Guide, 2004.
[45] A.M. Rugman, A. Verbeke, A perspective on regional and global strategies of multinational enterprises, Journal of International Business Studies 1 (2004) 3–19.
[46] J.H. Dunning, M. Fujita, N. Yakova, Some macro-data on the regionalisation/globalisation debate: a comment on the Rugman/Verbeke analysis, Journal of International Business Studies 38 (2007) 177–199.
[47] D. Somaya, Patent strategy and management: an integrative review and research agenda, Journal of Management 38 (2012) 1084–1114.
[48] Y. Kazakov, M. Krötzsch, F. Simancík, Concurrent classification of EL ontologies, Lecture Notes in Computer Science 7031 (2011) 305–320.
[49] Y.B. Yang, Y.F. Li, S. Krishnaswamy, Predicting reasoning performance using ontology metrics, Lecture Notes in Computer Science 7649 (2012) 198–214.