

A new patent processing suite for academic and research purposes



David Reymond^{a,*}, Luc Quoniam^b

^a Université de Toulon, I3M, EA 3820, 83957 La Garde, France

^b Aix-Marseille Université, IRSIC (EA 4262), 21, rue Virgile Marron, 13392 Marseille Cedex 05, France

ARTICLE INFO

Article history:

Received 3 May 2016

Received in revised form

5 October 2016

Accepted 5 October 2016

Available online 13 October 2016

Keywords:

Patent retrieval

Patent mining

Visualization

Education

Corpus creation

EspaceNet

ABSTRACT

Patent databases are a counterpart of a technical encyclopedia providing a valuable informational source from patent documents. Patents apply to a wide variety of uses, and this is expanding. The increasing kinds of users of patent documents result in a lack of educational programs regarding this matter. This paper describes Patent2Net (P2N) a patentinformatic suite whose purpose is to fill the lack in the academic world (education and research) of a tool to use with students (STEM, Masters, PhD), by valuation services and for defining corpuses for research in general patent analysis, specifically on textual content. P2N is a free open-source modular, scalable, customizable and derivable tool, written in the python language. We present here the main functions of the tool and the technical aspects after discussing the skills to be reached by students for state of the art patent analysis.

© 2016 Elsevier Ltd. All rights reserved.

1. Patents, users, uses and perspectives

Information about patent information has, so far, mostly remained a field for intellectual property (IP) experts working in industrial or legal frameworks. Patent-related documents are neglected for educational purposes [1] or are even absent from the bibliographic search processes of academic researchers [2]. It arises that patent analysis may answer many questions beyond those of mere technological intelligence [3]. In the field of scientometrics, patent analysis or technometry [4] provides indicators for assessing, for example, the development of research groups and technology fields [5] and helps develop research policies. Patentinformatic tools, which may be defined as software used to process data about patents, assist those seeking information about patents by combining four main core fields: bibliographic citation [6], content analysis [7], networks [8] and statistical analysis. In the following, we discuss the main characteristics of patent search and analysis in order to underline the main functions for the study of patent information. We also point out the growth of users and uses that highlights the need for educational programs besides research activity on patent mediation. We start by reviewing basic principles about patents.

2. The patent document as a source of information

A patent plays a strong part in information dissemination. A patent is a structured document, its prosecution status is recorded in a database and experts who provide opinions and classification metadata examine it. Each stage of the preparation and prosecution of a patent application generates information. At the first stage of its preparation (Fig. 1: Life cycle of a patent idea - from [8], p. 9), the pre-filing stage document must describe the invention with a description, drawings and possibly a bibliographic search for references that may be representative of the field as so-called background art. The search for background art may include, for example, a search into patent databases using a patent class search. Intellectual property offices label each patent application with a few classification numbers corresponding to the patent's technical field. To facilitate patent search, various patent offices, such as the European Patent Office (EPO) and the United States Patent and Trademark Office (USPTO), are phasing in a *Cooperative Patent Classification* system. Full text search, such as using keywords and their proximity, is a relatively new function [9], p. 28 in the patent retrieval process.

A patent application may include a background section describing the problem to be solved and the current state of the art, academic literature or other patent publications. A patent application includes a set of claims that are written in legal terminology and style [8], p. 9.

* Corresponding author.

E-mail address: david.reymond@univ-tln.fr (D. Reymond).

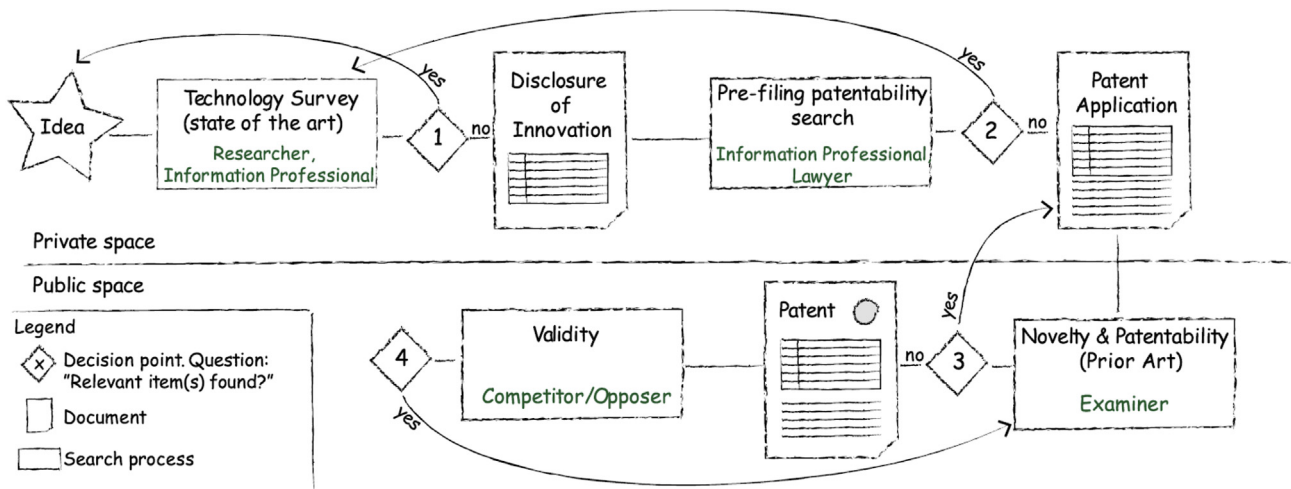


Fig. 1. Life cycle of a patent idea (from Ref. [9], p. 9).

Following the filing of a patent application document with a patent office and the payment of filing and examination fees, the patent application undergoes an examination process, and it usually takes from three to eight years to obtain granting of the patent. Once the patent granted, the patent holder may license the technology described in the claims or exclude others from using it for up to 20 years in selected countries from the date the patent was filed. The examination process results publishing a granted patent, assigned classification labels and related to further background art that may have been found by the examiner. Most of the examination process is a set of communications between the examiner and inventor that usually results in adjusting the claims to limit their coverage in response to prior art found by the examiner, thereby more precisely defining the scope of the invention. During the examination process anybody may submit prior art to the patent office to facilitate free use of the patent subject. After granting, a patent may be challenged by a competitor aiming to invalidate or narrow the scope of the patent in a litigation process to the attention of the patent office.

Hence, most of the information found in patent documents is free to use to develop and market processes or products. However, patents are, in general, a misused information source [10] [1]. denoted the need to start educational programs in patent analysis for science and engineering students in France. This idea was followed by Ref. [11]: 'There is a need both for a formal education program to sit alongside the experience gained in practice, and an agreed standard to provide formal recognition of the knowledge, skills and experience vested in the competent patent information professional'. On the other hand, for scientific research [2], noted the need to use patents as a research object in itself in order to, for example, determine the state of the art in a domain, to elaborate research policy strategy and for library services [12]. All the authors have denoted the need for patentinformatic tools to fill the gap in academic research and education as usual software and online services in this domain may be too expensive for educational and academic purposes. See Refs. [13] and [14] for a comparison of several patent analysis tools. We note that the specific value added by private databases is marginal for educational purposes as a key skill is to address a database correctly and not a specific skill for which complementary skills would be obtained 'at job experience'.

3. A patent database as a technical encyclopedia

Patent database search capabilities may differ in terms of geographic and historical coverage, the type of documents

available, the descriptive fields of the documents that are accessible and searchable (titles, abstracts, classification, date, etc.) and their contents (description, claims or other elements). In addition, many databases offer various tools designed to facilitate the retrieval of relevant search results as well as the visualization and analysis of these results. Among the many patent databases that are freely accessible on the web (see Ref. [15], EspaceNet is a free patent search service offered by the European Patent Office (EPO). EspaceNet permits full text search for some patent documents in English, German and French and offers search possibilities through bibliographic data (INPADOCDB) and legal status data sets (INPADOC). The 'worldwide' database combines over 90 million patent documents dating back to 1836 in 89 databases managed by over 50 patent authorities. EspaceNet lacks tools for data analysis, but, in 2006, the European Patent Office released an *Application Programming Interface* (see the EPO API Console for their database called *Open Patent Services* (OPS)) [16]. API enables the development of software to freely access the database under conditions of fair-robot and fair loading of servers up to a weekly 2.5 GB limit.

The EspaceNet worldwide database has been used previously to examine patents in several fields, such as genetics and biology [17], the globalization of knowledge [18] and nanotechnology [19]) [16]. suggests using the OPS API 'to extend usages'.

Due to the editing process of patent documents from filing to issuance, a patent application (prior art examination in Fig. 1) must contain novel and unique claims to be granted. Information contained in patents is typically more detailed and exhaustive than that found in scientific papers [20]. The language of patents may be highly specialized and include technical words not found in everyday language [21]. However, patents are also legal documents that are often written in 'patent jargon'; the knowledge they contain may be difficult to utilize and understand [22]. These points make us view the EspaceNet database as a technical encyclopedia [2] the use of which is not obvious.

4. The spread of patent uses: the tasks

Traditionally, patent information tasks involve searching patents and analyzing them and monitoring new publications [14]. Monitoring and analysis of new publications under time constraints is beyond the scope of this article.

Patent search (or patent retrieval) is a subdomain of Information Retrieval. Patent search is a general term that covers different types of search processes, such as a technology survey, prior art search, freedom to operate search, validity analysis and patent portfolio

Download English Version:

<https://daneshyari.com/en/article/6453173>

Download Persian Version:

<https://daneshyari.com/article/6453173>

[Daneshyari.com](https://daneshyari.com)