

Increasing the Discovery and Use of Non-Patent Literature (NPL): Scientific Publications in Patent Examination

Fernando LOIZIDES^{a,1}, Barrou DIALLO^b, Andrew POLLARD^a, Aekaterini MAVRI^c

^a*University of Wolverhampton*

^b*Independent Author*

^c*Cyprus University of Technology*

Abstract. In this work we present findings on non-patent literature use, and specifically scientific publications such as academic articles. We interview patent examiners and observe their prior art searching in order to provide insights into the perceived usage of non-patent literature and produce high level requirements for advancing non-patent literature search tools.

Keywords. Patents, Patent Examination, Non-Patent Literature, NPL, Scientific Literature, Academic Publications, Intellectual Property

1. Introduction and Motivation

Patents give the sole rights to an invention to an individual or company on a unique novel item. Patent applications have increased dramatically yearly in the past few decades [1, 2, 3, 17]. Every year, billions are spent in lawsuits over patent infringement and legal battles, making these publications an expensive commodity [4]. In order for a patent to be accepted and published, a patent application is filed and an examination occurs in which prior art is sought, in order to dispute the novelty of the application. A patent examiner triages [5] through large amounts of publications in order to determine the uniqueness of the invention and make a decision as to whether to publish, require amendments and resubmit, or to reject the patent. Being able to search through the previous publications in this scenario is vital so as not to grant a patent to an already existing invention. Therefore, the search on existing publications needs to be exhaustive.

The main search tools and repository used by patent examiners to search for prior art is currently local databases available to the patent examiners. These can be heavily focused on patent applications and previously awarded patents. There are also however further resources that also classify as prior art. Examples of other types of documents include academic articles and also internet based literature. The last two categories are also deemed as extremely relevant and useful to patent examination [15]. Indeed “the knowledge generated by academic scientists has been deemed one of the most crucial ingredients for technological progress and economic growth” [14] and therefore is vital

¹ Corresponding Author's e-mail address: Fernando.loizides@wlv.ac.uk

to be included in the searching for prior art. It is therefore “irrelevant if prior art is disclosed in patent literature or alternatively in non-patent literature (NPL), including scientific/technical journals and all other kind of content available e.g. on the internet and elsewhere. For in depth analysis of prior art comparable search efforts should therefore be undertaken for non-patent literature as for patent literature” [11]. In fact, many times instead of merely complimenting the search, other types of published work is sufficient to disprove the novelty or prove the existence of already existing prior art on its own [12].

In this work we present and discuss requirements that exist among patent examiners regarding the search and use of existing scientific publications (prior art) and underlying reasons for these. We define scientific publications in this work to mean peer reviewed academic papers such as those found in conference proceedings or in journals. We bring together work from different existing studies within literature and through ongoing user studies conducted within the European Patent Office in The Hague, Netherlands with interviews from patent. The work presented in this paper provides the foundation for subsequent research and development and examines the factors which contribute to the perceived value of these publications such as their findability and accessibility (for example the lack of open access). The value of our work can be seen in three main aims: a) raising awareness of the importance of these publications within patent literature b) eliciting requirements reasons for the discovery and utilization of scientific literature within patent examinations and c) inform technological solutions which may be created or improved upon in future.

2. Background

There is evidence that the use of academic articles within patent examinations and prior art searching is both useful and being sought out within the patent examination domain. For example even in the early 1990's access to use internet resources were made available to the patent examiners, even to a limited degree [6, 7]. In a 2001 internal report 1/30 search reports by the European Patent Office had reference to an internet document. 1/5 search reports within the European Patent Office cite Non-Patent Literature [19]. Guidelines for citing documents from the Internet and their importance were proposed [8]. The European Patent Office has therefore encouraged the use of and included scientific articles and resources within internal publication databases [9]. The process of using cross-organizational resources and digital libraries in order to “overcome the current information deficit and to fulfill the information need of the experts in the innovation-process” was also accelerated and investigated [10]. The prior work and evidence point to a need for a more thorough and up to date investigation and towards a more formal user-requirements based testing and analysis, before setting up any of the mentioned infrastructures for accessing publications. Authors, readers and publishers have now started to move towards an open access paradigm and encourage the availability of academic articles to the general public. We recognize user needs that need to contribute to the design of a framework and tools for patent application analysis with regards to prior art searching, an area that is receiving exponential need and growth [16].

3. Methodology and Findings

We identify initial high level points to assist both in creating suitable repositories as well as tools for searching and presenting information. The results presented in this section originate from previous published work as well as interview sessions from patent examiners at the European Patent Office. Specifically, 18 examiners were interviewed in depth while a further 14 examiners were observed while undertaking a prior art search followed by semi-structured interview. The examiners varied in experience, from 2 to 31 years (AVG 16 years - STDEV 8). The examiners fell under categories A-G (<http://www.wipo.int/classifications/ipc/en/> - accessed April 2017). The average amount of prior art searches conducted per year were 70 with a standard deviation of 30. All patent examiners reported this number increasing yearly, indicating that there is a need for further support for faster prior art searches in the same amount of time. Most examiners had extensive experience with scientific publications, usually from their university studies (such as PhD work). All examiners were familiar with scientific publications.

We began by questioning the examiners of their *perceived importance* of non-patent literature, and specifically scientific literature, to the prior art search process. All but one of the examiners saw importance and relevance to searching through scientific documents. The main value was seen in applied research. The examiners all agreed that some fields relate less to academically published material than others. Examples brought forward of research which is highly relevant to scientific publishing regarding their likely relevance in prior art include Informatics areas such as Bioinformatics, Medical informatics and Machine Learning. Examples where scientific literature is perceived to have less of an impact in prior art search include areas such as the furniture and shoe industry, which are more design oriented. Upon questioning the examiners, if they think examinations include NPL responses varied. NPL searching was reported as taking place in specific fields (consistent with the feedback on whether NPL is considered relevant) and on occasion to a limited degree. The question was put forward to the patent examiner which does not use NPL of whether they “think that if they searched for academic literature the outcomes of the examinations may be different” the examiner answered that “to a small extend maybe, some prior art might be cited even if it is less relevant”. All but one examiner reported performing NPL search during their prior art searching under 90% of the time. The one remaining examiner reported searching for NPL on every prior art search conducted. In terms of perceived success half the examiners reported 1/3 or less of the times being able to successfully find relevant scientific publications in their prior art searching, 2 of our pool of examiners did not answer. The remaining examiners reported that “most searches can have relevant literature”.

3.1. Searching non-patent full text.

Some platforms allow for full text searching. This is however; still weak compared to functionalities in patent full text databases [15]. This was identified by the examiners to be “cumbersome to get the full article sometimes (need to order the article), even if it became much easier in the last years”. The search from a search engine is also often done on the abstract or keywords rather than the full text with the user needing to then

manually triage through the full text in order to locate relevant sections. There is a need to allow a full text search directly from the search engine as well as improve tools for the information seeker to investigate the full text internally efficiently.

3.2. Central searching location vs multiple repositories.

One of the current issues that examiners and information seekers face is “to combine several publication search platforms that must be searched separately, with different search interfaces [15]. Search tools should combine different sources and repositories using a familiar common searching interface (see also [11]). Currently, half the examiners reported using mainly external tools in order to search for NPL rather than internal tools. The main reported sources by the pool of participants included Google Scholar, generic Google searching, YouTube and Orchid.

3.3. Different document types.

There are currently different ways to access articles (PDF, HTML, Word, Plain text) which makes the discovery and use of the articles challenging. A common format upon submission by the authors or tools that convert to a standard such as XML should be made available. This would allow for the easier searching of within document material by search engines as well as creating tools for investigating document content by the information seeker. Where there is no automatic XML type conversion within the publisher, an automated medium tool can allow for on-the-fly XML conversion of the documents. It is also worth noting that there is often information associated with the scientific publications which although relevant and useful, may not always be readily available. As one examiner noted: “We miss often the possibility to search conference materials (powerpoint presentation, handbook of abstracts) that are brought up to our attention later by third party or during oppositions”.

3.4. Provide familiar and advanced search options.

This principle is one that holds true as a generic usability principle regarding user interfaces. Searching should provide options to allow the user the freedom to customize the search and be more advanced in how they narrow down what they need. Examples include Boolean logic, keyword highlighting tools based on search and image extraction. Examiners reported on using external (to the patent office) tools in order to search for NPL such as scientific documents. As one of the participant reported, “different tools” require “different query languages”. Using these tools produced some difficulties of familiarity. For example, some tools are reported to not “prioritize the documents as you [an examiner] would”. External tools are also characterized as a “black box” whereas an examiner would like to know how the search is performed. There are often also problems with NPL searching tools such as the opportunity to structure the search in a logical manner rather than searching and losing the search result flow after another search query is entered. The most common search limitation that was reported for scientific publications, is that of “lack of standardized classification of documents”. Unlike patent literature that is classified to a detailed level, scientific publication classification varies. Different conferences and journals have their own classification systems which differ to each other, sometimes to a large extent. The other form of classification on scientific documents is a ‘keywords’

classification, which is often given by the authors and therefore may be even more diverse than the publisher's. It is therefore almost impossible to have a universal classification system implemented on scientific publications. This makes the ability to be exhaustive and the discoverability of the documents extremely difficult for the examiners. This limitation coupled with the fact that NPL search is "more time consuming compared to patent searching", can decrease "efficiency due to lack of time". This results in a 'satisficing' scenario [18] where examiners "draw the line somewhere in terms of time efforts" when a 'reasonable rather than exhaustive' search is made.

3.5. Specialized presentation and navigation

Examiners made a clear point on how "interfaces are not standardized". Tools for specialized searching should include an interface customized to assist in the rapid searching and assimilation of information. Currently, the databases and interfaces in existence for scholarly searches can rate less than adequate. Examples to address this issue is faster navigation to different 'important identified' sections of a document, image extraction and presentation and faster switching between different documents in a search. Another large requirement from the patent examiners are those of serendipity searching by navigation. This was defined as "starting from one document, to find all related documents to this starting document multidimensionality by filtering for author name and from backward and forward citations".

4. Summary and Future Work

We present work investigating the internal usage and perceived value of non-patent literature within prior art searching. From interviews with patent examiners we are able to produce high level guidelines to inform the creation of software to assist prior art searching for NPL. In future, we aim to present a prototype tool which is designed based on these guidelines and which will be able to elicit further detailed requirements while using a user-centered design approach to continue the development of the interface.

5. Acknowledgments and Disclaimer

We would like to thank the patent examiners for volunteering their free time and for their insights which made this research work possible. The opinions of the authors do not necessarily reflect the opinion of their employers.

References

- [1] <http://www.epo.org/about-us/annual-reports-statistics/statistics.html#filings> (accessed – April 2017)
- [2] Hingley, P., & Bas, S. (2009). Numbers and sizes of applicants at the European Patent Office. *World Patent Information*, 31(4), 285-298.
- [3] van Zeebroeck, N., Stevnsborg, N., De La Potterie, B. V. P., Guellec, D., & Archontopoulos, E. (2008). Patent inflation in Europe. *World Patent Information*, 30(1), 43-52.
- [4] List, J. (2008). Free patent databases come of age. *World Patent Information*, 30(3), 185-186.
- [5] Creswell, J. (2006). So small a town, so many patent suits. *New York Times*, 24.
- [6] Badi, R., Bae, S., Moore, J. M., Meintanis, K., Zacchi, A., Hsieh, H., ... & Marshall, C. C. (2006, January). Recognizing user interest and document value from reading and organizing activities in document triage. In *Proceedings of the 11th international conference on Intelligent user interfaces* (pp. 218-225). ACM.
- [7] Jonckheere C. The Internet for search project at the EPO. *World Patent Inform* 1997;19(2):93
- [8] Black T, Hayes G, Lintz P, Purcell AF. Use of the Internet in the USPTO for the patent examination process. *World Patent Inform* 1997;19(3):153.
- [9] Verhulst W. The impact of the Internet on prior-art searching in a patent environment. *World Patent Inform* 1997;19(2):95.
- [10] Archontopoulos, E. (2004). Prior art search tools on the Internet and legal status of the results: a European Patent Office perspective. *World Patent Information*, 26(2), 113-121.
- [11] Landwich, P., Vogel, T., Klas, C. P., & Hemmje, M. (2009). Supporting patent retrieval in the context of innovation-processes by means of information dialogue. *World Patent Information*, 31(4), 315-322.
- [12] Annies, M. (2009). Full-text prior art and chemical structure searching in e-journals and on the internet—A patent information professional's perspective. *World Patent Information*, 31(4), 278-284.
- [13] Bonino, D., Ciaramella, A., & Corno, F. (2010). Review of the state-of-the-art in patent information and forthcoming evolutions in intelligent patent informatics. *World Patent Information*, 32(1), 30-38.
- [14] Salampasis, M., & Hanbury, A. (2014). PerFedPat: An integrated federated system for patent search. *World Patent Information*, 38, 4-11.
- [15] Adams, S. (2012). Survey of PCT search reports and the importance of the internet as a source of non-patent literature. *World Patent Information*, 34(2), 112-123.
- [16] Baudour, F., & van de Kuilen, A. (2015). Evolution of the Patent Information World—Challenges of yesterday, today and tomorrow. *World Patent Information*, 40, 4-9.
- [17] Abbas, A., Zhang, L., & Khan, S. U. (2014). A literature review on the state-of-the-art in patent analysis. *World Patent Information*, 37, 3-13.
- [18] Winter, Sidney G. "Satisficing, selection, and the innovating remnant." *The Quarterly Journal of Economics* 85.2 (1971): 237-261.
- [19] http://www.wipo.int/edocs/mdocs/scit/en/scit_6/scit_6_www_31685.pdf (Accessed - 2017)