

## Role of patent analysis in corporate R&D



Cynthia Barcelon Yang\*

Innovation is the foundation of today's global competitive economy and technological advancement. Throughout the business process – from an idea to a product on the market – information, and specifically patent information, is the currency and key driver of the innovative process. Patents reveal new knowledge through disclosure of inventions and enable others to develop further technological innovations by avoiding duplicate R&D efforts. Through licensing transactions, patents enhance the technology-transfer processes and commercialization of new technologies [1]. This article describes the role of patent analysis in the innovative process within the context of a corporate R&D environment, focusing on the use of text-mining tools.

According to the WIPO Statistics Database [101], an increasing trend of worldwide patent information is becoming evident, especially among the top five patenting authorities – USPTO, EPO, SIPO, KIPO and JPO. Roughly one million patent applications are filed each year worldwide. The paradox today is that the exponential growth of patent information makes it very difficult to find the true state of the art and usable knowledge. The challenge is how to keep up with this flood of information. Spotting patterns and extracting useful and actionable information are even harder since much of the information in patents is unstructured text. Mining this wealth of information is a laborious and expensive task that requires meticulous, mostly manual reading and analysis of patent documents.

Furthermore, there are challenges resulting from industrial competition and a changing global IP landscape. The race to have the strongest and largest IP portfolio is increasingly evident in today's competitive markets. Patents are being more zealously acquired, vigorously asserted and aggressively enforced. Meticulous and vigilant efforts are required in analyzing results and finding all relevant patent information in order to avoid or mitigate these competitive risks. As emerging markets in Asia, Latin America and Eastern Europe are making inroads in strengthening and enforcing their patent laws and practices, how to keep up and find patents from these emerging countries and translate them into English or other languages are new challenges that must be overcome for effective patent analysis and IP protection.

Patent analysis plays a significant and essential role in the innovative process by delivering timely and value-added patent information to support three key

\*Bristol Myers Squibb, PO Box 5400,  
Princeton, NJ 08543-5400, USA  
E-mail: [cynthia.yang@bms.com](mailto:cynthia.yang@bms.com)

**Keywords:** patent analysis • patent information • patent law • technology innovations  
• text-mining tools

areas in corporate R&D. The first area is to support the scientific R&D, for example, by providing state-of-the-art analysis on a particular technology area. The second area is to support IP procurement and protection. Patent attorneys need analyzed patent information to formulate legal opinions on freedom to operate issues, determination of patent validity and advice on sourcing and procurement, as well as in support of due diligence work. The third area is to support business development and strategic transactions. In business-development work, patent information is important for analysis of competitors, for understanding a particular technology area, determination of patent enforceability, and identification of potential alliance partners for in-licensing and investment activities, and for drafting claims for out-licensing and divestment of patents.

Manual human analysis, which we have relied upon for the past 50 years or more, is no longer a sustainable strategy. More innovative approaches are required. With advances in computing and linguistic technologies, new text-mining and visualization tools have emerged in the marketplace, especially in the last decade. These tools present an opportunity to assist and facilitate the patent search and analysis process.

**“ Patents are being more zealously acquired, vigorously asserted and aggressively enforced. ”**

The majority of text-mining tools that have emerged in the past decade are macro-analytics types of tools that work well with structured data or metadata. These are based largely on statistical algorithms or frequency of words appearing in the text, with no insight or inference to the meaning of the words within the textual context. These tools primarily perform two types of analyses – technology assessment and company profiling. Technology-assessment analysis provides high-level visualization of a large dataset of patent information, and a quick overview of the various segments in a particular technology domain to support decision-making (e.g., finding key organizations, key researchers, and patenting trends in that technology or related technologies). Company profiling, on the other hand, consists of visual representations of the patent assets of a specific company and provides insights into the company in terms of patenting trends, geographic protections, top inventors, its research focus and collaboration partners.

#### Case studies

A case study on hepatitis C virus (HCV) inhibitors illustrates the technology assessment analysis [2]. Typical questions asked in a technology assessment include:

‘what are the top companies patenting in this technology area, in this case, HCV inhibitors in the last 8 or 10 years?’; ‘which companies and/or inventors are closely collaborating?’; ‘who are the prolific inventors whom we might potentially collaborate with or hire?’. In this case study, the VantagePoint™ tool was used with data from MicroPatent™ consisting of full-text patents from the PTO, WIPO and EPO. Keywords related to HCV inhibitors from 2001 to 2007 were searched, resulting in 540 retrieved patents, based on 299 unique patent families or inventions. A chart showing the top 15 companies and their patenting trends in the HCV inhibitor area was produced. A correlation map showing the interactions among the various assignees (represented in the map as circles with varying diameters) based on whether documents have common inventors was also plotted. In VantagePoint, mousing-over each circle brings up a list of the major inventors associated with that particular company. Assignees connected by a solid line indicate a stronger degree of collaboration on certain aspects of their research.

A second case study involved a company profile of Kosan Biosciences [2]. In this case study, a patent assignee search on Kosan was conducted using Derwent World Patent Index and Chemical Abstracts Services as database sources. Approximately 120 patent records were retrieved. An inventor auto-correlation map that indicates who is working with whom shows the different research teams working together in Kosan. A 3D matrix shows Kosan’s research focus in terms of mechanism of action (HSP90 inhibitors) and also possibly what programs were likely abandoned. A huge spike on cancer treatment is evident from a 3D-matrix plot on Kosan’s research focus in terms of utility or use. Another way to analyze and view this company patent profile is using research landscape tools, such as STN Anavist™. The most frequently occurring concepts of HSP-90 inhibitor for cancer treatment are displayed as a cluster. Clicking on this cluster highlights in a specified color (e.g., yellow) those document titles with related key researchers, publication year and clustering concepts.

**“ The majority of text-mining tools that have emerged in the past decade are macro-analytics types of tools that work well with structured data or metadata. ”**

While there have been advances in macro-analytics tools, there are still limitations and further improvements are required. The main limitation is that these tools typically work well only for structured text or metadata, such as patent number or patent assignee. The most significant enhancement will be applica-

tion of micro-analytics technologies to unstructured text, especially for finding the needle in the haystack. Specific questions such as: ‘what compounds bind to protein X?’; ‘what are the alternative disease(s) our compounds would be effective in?’ or ‘what dosages of compound B cause an effective biological response or adverse reactions?’ require inferring relationships between entities. Aided by ontologies, micro-analytics tools (e.g., Linguamatics™) utilize linguistic and semantic-based algorithms to extract relationships between entities and interpret the meaning of unstructured text and highlight them as ‘assertions’ in the results output, rather than just retrieving documents that one has to read to obtain the answer. They go beyond the simple information retrieval approaches commonly used, such as Google or Bing.

## Conclusion

To paraphrase Kesting and Woods’ statement [3], the role of a patent analyst as an intelligence professional is not to gather information to pass on to decision makers, but to design and analyze the results into ‘knowledge blocks’ that enable decision makers to solve their own puzzle. In the application of patent-analytics tools as illustrated here in the two case studies, the analyst selects the types of analysis and visualization options most appropriate to the business question and to the dataset of interest, based on the patent analyst’s scientific expertise, business knowledge, understanding of client’s needs and knowledge of tools and databases. The patent analyst collaborates interactively with the clients to refine the analysis criteria and output. As patent information is a crucial element in the innovative process for business success, especially in the pharmaceutical industry, the patent analyst plays a critical role

in ensuring thoroughness, timeliness and accuracy of deliverables. Simply ‘Googling it’ is not a sustainable research strategy. Value-added, accurate and comprehensive patent content combined with innovative technology tools are required to address the increasing volume and complexity of patent information and deliver reliable, comprehensive patent intelligence to stakeholders.

**“...the analyst selects the types of analysis and visualization options most ... appropriate to the business question and to the dataset of interest...”**

## Acknowledgements

*This article was developed from a presentation given by the author at the Patent Statistics for Decision Makers 2011 Conference, US PTO, Alexandria, VA, 16-17 November 2011. The author would like to thank the BMS Patent Analysis Group, specifically YY Yang, L Akers, T Klose, S Pavlek and A Goshorn, as well as Search Technology, Inc., Chemical Abstracts Service and Linguamatics for their contributions and technical support to the work described here.*

## Financial & competing interests disclosure

*The author is an employee of Bristol-Myers Squibb. She is also a member of the Board of Directors of the Patent Information Users Group, Inc. (PIUG), an IP Advisory Board Member of the Information Retrieval Facility (IRF), a member of the Scientific and Technical Information Network (STN) Advisory Council and the Patent Documentation Group (PDG). The author has no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed. No writing assistance was utilized in the production of this manuscript.*

## References

- Guellec D, van Pottelsberghe B. *The Economics of the European Patent System: IP Policy for Innovation and Competition*. Oxford University Press, Inc., NY, USA (2007).
- Yang YY, Akers L, Yang CB, Klose T, Pavlek S. Enhancing patent landscape analysis with visualization output. *World Patent Inform.* 32, 203–220 (2010).
- Kesting W, Woods K. Experiential modeling: innovation opportunity for competitive intelligence professionals. *Competitive Intelligence Rev.* 7, 57–68 (1996).
- Website
- WIPO IP Facts and Figures, 2011. [www.wipo.int/export/sites/www/ipstats/en/statistics/patents/pdf/wipo\\_pub\\_943\\_2011.pdf](http://www.wipo.int/export/sites/www/ipstats/en/statistics/patents/pdf/wipo_pub_943_2011.pdf) (Accessed 10 December 2011)